

University of Groningen

Reliability of assessment of adherence to an antimicrobial treatment guideline

Mol, Peter G. M. ; Gans, Rijk O.B. ; Panday, PVN; Degener, John; Laseur, M; Ruskamp, Flora

Published in:
Journal of Hospital Infection

DOI:
[10.1016/j.jhin.2004.11.022](https://doi.org/10.1016/j.jhin.2004.11.022)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2005

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Mol, P. G. M., Gans, R. O. B., Panday, P. V. N., Degener, J. E., Laseur, M., & Haaijer-Ruskamp, F. M. (2005). Reliability of assessment of adherence to an antimicrobial treatment guideline. *Journal of Hospital Infection*, 60(4), 321-328. DOI: 10.1016/j.jhin.2004.11.022

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Reliability of assessment of adherence to an antimicrobial treatment guideline

P.G.M. Mol^{a,b,*}, R.O.B. Gans^d, P.V. Nannan Panday^b, J.E. Degener^c,
M. Laseur^b, F.M. Haaijer-Ruskamp^a

^aDepartment of Clinical Pharmacology, University Medical Center Groningen, The Netherlands

^bPharmacy Department, University Medical Center Groningen, The Netherlands

^cDepartment of Clinical Microbiology, University Medical Center Groningen, The Netherlands

^dDepartment of Internal Medicine, University Medical Center Groningen, The Netherlands

Received 18 October 2002; accepted 26 November 2004

Available online 20 April 2005

KEYWORDS

Antibiotic therapeutic use; Drug utilization review methods; Observer variation; Reproducibility of results

Summary Assessment procedures for adherence to a guideline must be reliable and credible. The aim of this study was to explore the reliability of assessment of adherence, taking account of the professional backgrounds of the observers. A secondary analysis explored the impact of case characteristics on assessment. Six observers (two hospital pharmacists, two internists and two clinical microbiologists) assessed a random sample of 22 prescriptions made to infectious disease cases admitted to a department of internal medicine between February and August 2001. Agreement between observers with regard to adherence of these prescriptions to guideline recommendations concerning drug choice, duration of treatment, dosage and route of administration was measured using Cohen's kappa. Case characteristics were compared between cases where observers agreed and disagreed with two-sided Fisher's exact test. Agreement between all professionals was moderate for drug choice (0.59), fair for duration of therapy (0.36), moderate for dosage (0.48), and fair for route of administration (0.37). Agreement on drug choice was good within (0.75 and 0.83) and between (0.74) the internists and the hospital pharmacists, but was less within (0.31) the clinical microbiologists and between the clinical microbiologists and the internists (0.44) and the hospital pharmacists (0.42). Within the clinical microbiologists, agreement was good for dosage (0.79) and route of administration (0.66). There was frequent disagreement between observers regarding cases with combination therapy and non-immunocompromised patients. Despite the small number of cases,

* Corresponding author. Address: Department of Clinical Pharmacology, University Medical Center Groningen, P.O. Box 196, 9700 AD Groningen, The Netherlands. Tel.: +31 50 3638313; fax.: +31 50 3632812.

E-mail address: p.mol@med.umcg.nl

our results suggest that internists and hospital pharmacists can reliably be used to assess adherence for drug choice. The level of agreement seems to be affected by combination therapy and the immune status of the patient. © 2005 The Hospital Infection Society. Published by Elsevier Ltd. All rights reserved.

Introduction

The appropriate use of antimicrobial agents is considered to be a major factor contributing to increasing bacterial resistance for commonly used antimicrobials. Prudent use might curb development of new resistant bacteria and reduce prevailing antimicrobial resistance.¹⁻³ Programmes to optimize antimicrobial use have reduced the cost and volume of therapy while optimizing care.⁴⁻⁶ Adherence to hospital antibiotic guidelines is a commonly used primary outcome measure in studies of antimicrobial treatment patterns. In some of these studies, an infectious disease specialist reviewed medical charts to assess adherence to a guideline.⁷⁻¹² In other studies, junior clinical or hospital pharmacists, internists or residents, and clinical microbiologists reviewed prescribing appropriateness or adherence.^{4-6,13-20} However, low to moderate agreement was found by some authors in assessments.^{10,17} Satisfactory levels of agreement were only found between professionals of the same professional background, such as between hospital pharmacists, using the medication appropriateness index, an assessment tool for appropriate prescribing.²¹

Any assessment method needs to be reliable and credible for the target group being assessed. In studies where various professionals review prescribing by, for example, internists, it is relevant that agreement exists between observers of different professional backgrounds. The aim of this study was to evaluate the reliability of the assessment of adherence to guideline recommendations for antimicrobial therapy, taking into account the different professional backgrounds of the assessors. In a secondary analysis, we explored the influence of case characteristics on assessment.

Methods

The guideline referred to in this study was locally developed for treatment of various infectious diseases. Its recommendations included preferred drugs, duration of therapy, dosage, and route of administration. Recommendations were based on

international and national treatment guidelines and local resistance patterns. The antimicrobial treatment guideline was developed by a hospital antimicrobial use committee composed of clinical microbiologists, hospital pharmacists and representatives of various medical specialties. Since no universally accepted standard exists for adherence to antimicrobial treatment guideline recommendations, reliability was studied by interobserver agreement.

Twenty-two patients were selected at random from a database of patients; the 'source population'. The source population consisted of patients who were admitted to the department of internal medicine between February and July 2001 and who received an antimicrobial prescription for an indication covered by the hospital guideline. Data collected for these patients included: (a) patient characteristics including sex, age, drug allergies and co-morbidity, short medical history, and co-medication; (b) disease characteristics, e.g. temperature and indication for antimicrobial use as recorded by the ward doctor; and (c) laboratory findings, e.g. culture results, and liver and kidney function tests.

Six professionals (two internists, two hospital pharmacists and two clinical microbiologists) independently assessed the adherence of antimicrobial prescriptions to the hospital guideline recommendations. These six observers received instruction by letter and face-to-face from the main researcher (PGMM) on how to assess each prescription, as well as a copy of the prevailing hospital antimicrobial treatment guideline.²² The assessment was based on an algorithm designed for assessing antimicrobial prescribing by Kunin *et al.*²³ and adapted by Gyssens *et al.*²⁴ (Figure 1).

The unit of analysis was an antimicrobial prescription item (one drug). For each patient receiving antimicrobial treatment, often consisting of more than one prescription, one antimicrobial agent was selected at random (using a die) for analysis. Observers received a complete description of such a patient case, from admission to discharge, including all prescribed antimicrobial agents. The observers were asked explicitly to assess only the selected prescription for adherence to the

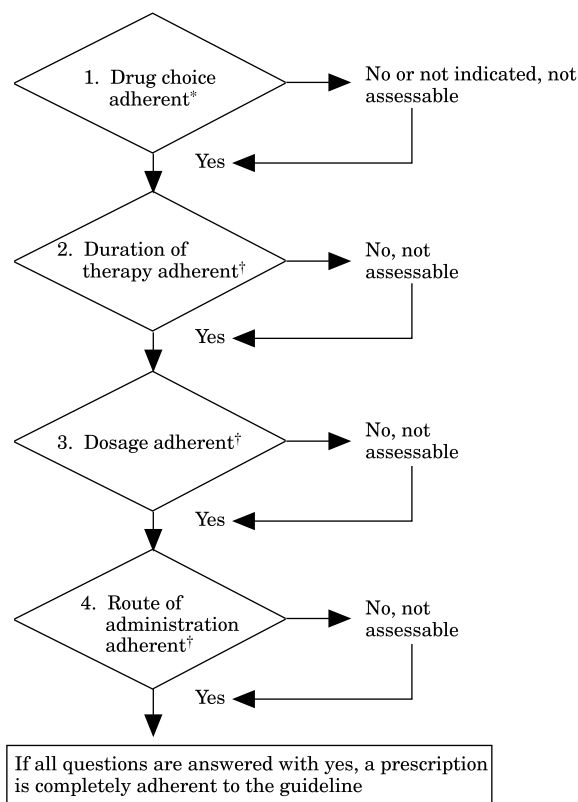


Figure 1 Assessment criteria. Adherence of an antimicrobial prescription with guideline recommendations.

guideline based upon the four assessment criteria: (a) drug choice; (b) duration of therapy; (c) dosage; and (d) route of administration. They were asked to take into consideration possible concurrently administered antibiotics. Cases could be assessed as adherent to the guideline, non-adherent or non-assessable (Figure 1). The observers were encouraged to give further comments on the overall treatment of a specific patient. Results were discussed with the observers after initial analysis in order to explore possible reasons for low agreement.

To examine the importance of case characteristics, three characteristics were selected. Firstly, combination therapy, defined as one or more antimicrobials given concurrently, which may influence agreement of assessment. For example, in combination therapy, one of the agents prescribed might be superfluous, but without explanation by the guideline regarding which is the redundant drug. Secondly, the immune status of the patient, since in immunocompromised patients, the preferred antimicrobial choice might differ for some indications from the standard drug treatment. Thirdly, culture-guided therapy, as opposed to empirical therapy, since it involves additional interpretation of culture laboratory findings. The

guideline's main recommendation for such cases was to use the most narrow-spectrum antimicrobial to which the cultured pathogen is sensitive, without explicitly stating what specific drug to use. The effect of case characteristics on agreement of assessment between professionals was studied for those specialists that were in good agreement in order to minimize the impact of interobserver variability. Cases were characterized as an 'agreement case' when all the observers agreed on the assessment of adherence on drug choice, and as a 'disagreement case' when all the observers did not agree.

Analysis

Characteristics of sampled cases were compared with the source population for categorical data with Yates' corrected Chi-squared test or two-sided Fisher's exact test when expected values in cells were less than five. Continuous variables that were not normally distributed were tested with the Mann-Whitney *U*-test. Analysis of agreement was by Cohen's kappa, average kappa for pairwise analysed pairs of observers and proportional agreement.²⁵⁻²⁷ A specifically designed software program, AGREE[®] version 7 (ProGAMMA, The Netherlands) for calculating agreement indices, was used to calculate kappa values. Kappa is the more robust of the two measures of agreement as it corrects for chance agreement. However, interpretation of kappa is influenced by unbalanced prevalences of judgments in each category of the contingency table, i.e. proportion of adherent, non-adherent and non-assessable cases.²⁸ This leads to differences in the chance expected frequencies. Proportional agreement was defined as the average percentage of observers agreeing with the modal assessment in each case. For both measures, agreement between assessors could be for adherent, non-adherent or non-assessable therapy. The advantage of proportional agreement is that it is a more direct and hence easier to understand measure of agreement.

Case characteristics were compared between 'agreement' and 'disagreement' cases with two-sided Fisher's exact tests.

We used the interpretation of kappa values as proposed by Altman, who distinguished five levels of agreement: poor (less than 0.20), fair (0.21-0.40), moderate (0.41-0.60), good (0.61-0.80), and very good (0.81-1.00).²⁵ For purposes of power calculation, we considered agreement to be optimal at kappa of 0.8 or higher and unacceptable at kappa of less than 0.4. A sample size of 18 cases had

a power of 80% and $\alpha=0.05$ to detect these levels of agreement, with a distribution of 45%, 45% and 10% over the outcome categories adherent, non-adherent and non-assessable, respectively. We increased the minimum number of cases required by 25% to 22. Such small sample sizes are common in other studies probing agreement between multiple observers.^{29,30}

Results

Characteristics of the 22 randomly sampled patients reflected those of the source population. Fewer sampled patients presented with infections other than respiratory or urinary tract infections or septicaemia than the source population (Table I).

Agreement (average kappa) between observers

The average kappa for all six observers indicated moderate agreement for adherence of these prescriptions to guideline recommendations concerning drug choice, fair agreement on duration of therapy, moderate agreement for dosage, and fair agreement for route of administration.

Proportional agreement was high for drug choice (86%), lower (74%) for duration of therapy, high

(87%) for dosage, and intermediate (80%) for route of administration criterion (Table II).

Kappa values on dosage were relatively low because of the uneven distribution of cases judged to be adherent (median 17 of 22 cases per observer). Ratings of other assessment criteria were more evenly distributed over the outcome categories: adherent, non-adherent and non-assessable. The observers considered duration of therapy to be non-assessable for a median of 11 cases, due to insufficient transparency of the guideline. In another seven cases, they considered the route of administration to be non-assessable, either because the guideline was not clear about the preferred route of administration or the prescribed antimicrobial agent had no oral formulation and therefore the route of administration had to be intravenous (Table II).

Agreement and professional background

Agreement for adherence of these prescriptions to guideline recommendations concerning drug choice was good among the two hospital pharmacists (average kappa 0.82), meeting our target value of 0.8, and among the two internists (0.75); agreement among the two clinical microbiologists was only fair (0.36) and was below our 0.4 threshold for acceptable agreement (Table III). Agreement was good (0.74) between the hospital pharmacists and internists. Agreement between observers from a

Table I Patient and case characteristics

	Source population	Sample	<i>P</i> values
Patients (<i>N</i>)	258	22	
Male (%)	150 (58%)	11 (50%)	0.353 ^a
Age (years) ^b	62 (49-73)	67 (47-73)	0.522 ^c
Length of stay (days) ^b	17 (9-26)	11 (8-19)	0.066 ^c
Deaths	30 (12%)	3 (14%)	0.732 ^d
Antimicrobial prescriptions per patient ^b	2 (1-3)	1 (1-3)	0.674 ^c
Combination therapy	132 (25%)	5 (23%)	0.972 ^a
Type of treatment			
Empirical therapy	310 (60%)	14/22 (64%)	0.871 ^a
Documented therapy ^e	177 (34%)	6/22 (27%)	0.678 ^a
Prophylaxis	34 (7%)	2/22 (9%)	1.0 ^a
Main indications for a prescription			
Urinary tract infections	53 (10%)	5/22 (23%)	0.075 ^d
Respiratory tract infections	133 (26%)	6/22 (27%)	1.0 ^a
Septicaemia	61 (12%)	5/22 (23%)	0.174 ^d
Other	272 (52%)	6/22 (27%)	0.038 ^a

^a Yates' Chi-square test.

^b Median (interquartile range).

^c Mann-Whitney *U*-test.

^d Fisher's two-sided exact test.

^e Prescriptions based on results of microbiological investigations.

Table II Agreement between observers for the Gyssens/Kunin assessment criteria

Assessment criteria	Kappa ^a	Median (interquartile range) number of cases per observer assessed as			Proportional agreement ^b
		Adherent	Non-adherent	Non-assessable	
Drug choice	0.59 (SE 0.09) ^c	8.5 (7-10)	13.5 (12-15)	0 (0-0.25)	86% (SD 17%)
Duration of therapy	0.36 (SE 0.08)	6.5 (4-7.25)	5 (3.75-7)	11 (9.75-12)	74% (SD 18%)
Dosage	0.48 (SE 0.11)	17 (15-19)	4.5 (2.75-7)	0 (0-0.25)	87% (SD 19%)
Route of administration	0.37 (SE 0.09)	13.5 (12.75-14.5)	1.5 (1-3.25)	7 (4.25-7.25)	80% (SD 20%)

^a Kappa: overall kappa for six observers; pairwise analysed per pair of observers.

^b Proportional agreement = $\sum_i \max(X_{ij}/m) 100\%$ with m observers assessing n cases into j categories (e.g. adherent, non-adherent, non-assessable). X_{ij} = number of observers who assign case i into category j . $\max(X_{ij}/m)$ the maximum proportion of case i assigned to a single category.

^c The standard error for pairwise kappa was calculated by the AGREE software, according to the formula shown by Schouten HJA in *Statistica Neerlandica*, 1982, pages 56–57.²⁷

clinical microbiology background, compared with an internal medicine and hospital pharmacy background, was moderate (0.51 and 0.51, respectively). One of the clinical microbiologists never reached a kappa higher than 0.62 with any of the other observers for the drug choice criterion. When consulting him after initial data analysis, it became clear that he had deviated from the instructions. Proportional agreement was high (83%) among and between observers from different professional groups, with the exception of a lower value of 59% among the clinical microbiologists.

Agreement for the duration of therapy criterion within and between observers from different

professional backgrounds was only fair to moderate (0.26-0.49). Proportional agreement ranged from 55% to 80%.

Agreement on dosage was good among the clinical microbiologists (0.79), moderate among hospital pharmacists (0.42), and poor among internists (0.13). Agreement between observers from one professional group and those from observers with a different background ranged from fair to moderate (0.38-0.58). Proportional agreement was generally high, with values ranging from 82% to 90% among and between observers from all professional groups, although agreement among internists was lower at 68%.

Table III Agreement between observers from different professional backgrounds (kappa/proportional agreement)

	Hospital pharmacy	Clinical microbiology	Internal medicine
Drug choice			
Hospital pharmacy	0.82 (SE 0.12)/91%	0.51/84%	0.72/92%
Clinical microbiology		0.36 ^a (SE 0.15)/59%	0.51/83%
Internal medicine			0.72 (SE 0.15)/86%
Duration of therapy			
Hospital pharmacy	0.49 (SE 0.16)/68%	0.37/80%	0.43/77%
Clinical microbiology		0.27 (SE 0.15)/55%	0.26/75%
Internal medicine			0.30 (SE 0.16)/55%
Dosage			
Hospital pharmacy	0.42 (SE 0.24)/82%	0.58/90%	0.38/88%
Clinical microbiology		0.79 (SE 0.14)/86%	0.48/89%
Internal medicine			0.13 (SE 0.18)/68%
Route of administration			
Hospital pharmacy	0.30 (SE 0.14)/68%	0.42/84%	0.25/81%
Clinical microbiology		0.66 (SE 0.14)/86%	0.42/86%
Internal medicine			0.21 ^b (SE 0.22)/59%

SE, standard error; HP, hospital pharmacists; CM, clinical microbiologists.

^a Two observers from each professional group assessed prescribing adherence. Internists agree significantly less than CM ($P < 0.004$).

^b Significantly less agreement between CM than between HP ($P < 0.02$).

Agreement for the route of administration criterion was good (0.66) among clinical microbiologists, but only fair to moderate among and between observers from all other professional backgrounds. Proportional agreement ranged from 59% to 86%.

For the criteria of duration of therapy, dosage and route of administration, kappa never reached the target value of 0.8 (Table III).

Agreement and case characteristics

Table IV describes agreement between observers on drug choice in relation to case characteristics. The analysis was limited to hospital pharmacists and internists, where we found good interobserver agreement on drug choice. Of the 22 cases analysed, 17 had 100% agreement, and there was disagreement in five cases. Disagreement scored relatively high in cases with combination therapy. Combination vs empirical therapy did not lead to a different outcome. Observers agreed more on cases with immunocompromised patients than on cases that involved immunocompetent patients.

Agreement was defined as 100% of observers agreeing; when we defined agreement as three out of four observers agreeing, there was disagreement in only two out of 22 cases. Combination therapy remained a significant issue (Fisher's two-sided exact test: $P=0.04$), although other case characteristics, immune status of the patient (Fisher: $P=0.09$) and type of therapy (Fisher: $P=1.0$) had no impact on agreement.

Discussion

Our results indicate that hospital pharmacists and internists agree when assessing adherence to a guideline on the choice of the antimicrobial agent.

Other aspects of drug treatment showed low interobserver reliability. This low agreement was partly due to the lack of explicitness of the guideline recommendations. Only the clinical microbiologists were in good agreement with each other on the dosage and route of administration criteria. The clinical microbiologists seemed to regard various aspects of an infectious disease case differently from the other professionals included in the study. Our results support the hypothesis that agreement when assessing drug choice is influenced by case characteristics such as combination therapy, immune status and, to a lesser extent, whether antimicrobial therapy was either empirical or based on microbiological results.

While assessment was reliable for adherence of these prescriptions to guideline recommendations on drug choice, the other assessment criteria showed less agreement. Lack of explicit criteria on duration of therapy in the hospital guideline hampered the assessment of cases for this criterion. More specific guideline criteria were often not possible in view of the many factors that determine the appropriateness of continuing an antimicrobial therapy and the lack of hard evidence on the optimal duration of antimicrobial therapy. Although kappa was low (0.49) for the dosage criterion, the high proportional agreement of 87% indicated that agreement was acceptable. Kappa may have been too conservative for this criterion. The unbalanced proportion of judgements in the outcome categories had an impact on the chance expected frequencies that influence the kappa statistics, as described in the Methods section.^{25,28}

Assessment of route of administration suffered from both unbalanced distributions of adherent and non-adherent cases and from a larger number of non-assessable cases. Again, lack of explicit guideline recommendations with regard to the route of

Table IV Case characteristics and agreement among hospital pharmacists and internists

Characteristic		Agreement	Disagreement	Total	Fisher's exact test (two-sided)
Combination therapy	Yes	2	3	5	$P=0.06$
	No	15	2	17	
Type of therapy	Empirical	11	3	14	$P=0.61$, comparing cases with documented and empirical therapy ($N=20$)
	Documented ^a	4	2	6	
	Prophylaxis	2		2	
Immune status compromised ^b	Yes	14	1	15	$P=0.02$
	No	3	4	7	
Cases		17	5	22	

^a Prescriptions based on results of microbiological investigations.

^b Immune status of patient is compromised or probably compromised, either iatrogenic or due to underlying disease; e.g. diabetes, alcohol abuse, leukaemia.

administration might have been the reason. The clinical microbiologists seemed to interpret the more implicit directions, regarding appropriate dosages, of the guideline in the same way.

Overall, in cases where the guideline lacked explicit recommendations, clinical microbiologists might be the most suitable professionals to assess appropriateness, i.e. to give individual advice on route of administration and dosage. They were in good agreement in our study on these criteria, and in daily medical practice, they were often consulted for decisions on whether a patient could safely be switched from intravenous to oral therapy, optimal dosing decisions and whether to stop or continue antimicrobial treatment.

The guideline left too much room for interpretation, leading to disagreement.²⁸

In some studies, disagreement between observers was resolved by having a single observer, which ensured consistent analysis but carried the risk of consistent wrong assessments being made.^{10,14,31} Another approach described was co-review of each case, which was feasible in experimental situations but might be excessively labour intensive in continuously running programmes.

Different opinions of professionals in a study investigating adherence to a guideline are detrimental to the reliability and credibility of the outcome measure. In clinical practice, however, different viewpoints of professionals on an infectious disease case are common and can be advantageous; for example, in antimicrobial management team meetings, optimal antimicrobial treatment could be decided by consensus, giving the attending doctors insight into alternative options.⁴ Hospital pharmacists and internists can be equally involved in assessing drug choice, depending on the local situation.

The impact of case characteristics on agreement was only explored for drug choice because overall agreement on assessment was most reliable for this criterion. The analysis of the impact of case characteristics was limited to observers from hospital pharmacy and internal medicine, as they showed good agreement on drug choice. Our study suggests that case characteristics might be relevant for assessment of adherence, although no definite conclusions could be drawn because of the small sample size. Moreover, combination therapy was found to be an issue complicating assessment. Most published intervention and observational studies have limited themselves to a predefined focused research problem with a more homogenous group of indications or antimicrobial prescriptions.^{5,7,12-15,18} Our findings imply that results of such studies

cannot be extrapolated to general adherent or appropriate antimicrobial prescribing habits.

This study used a small but adequate sample size in view of the power to detect relevant levels of agreement and disagreement. Differences in assessment between observers from various professional backgrounds seemed to increase with less explicit criteria in the guideline. As only two observers were included from each professional group, these differences may also have been caused by differences between individuals rather than between professional groups. The small sample size was a limitation for drawing firm conclusions on the influence of case characteristics, but this was only a secondary research question. The differences in case mix found in the sample in comparison with the source population were not relevant for reliability of assessment.

In conclusion, any study assessing prescribing quality should assess the reliability of the assessments themselves. Our study showed that assessment of adherence of prescriptions to hospital guideline recommendations could reliably be done for the choice of the drug by either internists or hospital pharmacists. More explicit guidelines made assessment of adherence more reliable.

Acknowledgements

The authors thank W.J.M.J. Rutten, J. Schouten, N. Meessen and T.S. van der Werf for their help with case assessment.

References

1. Tenover FC. Development and spread of bacterial resistance to antimicrobial agents: an overview. *Clin Infect Dis* 2001; **33**(Suppl):S108–S115.
2. Gould IM. A review of the role of antibiotic policies in the control of antibiotic resistance. *J Antimicrob Chemother* 1999; **43**:459–465.
3. Seppala H, Klaukka T, Vuopio-Varkila J, *et al.* The effect of changes in the consumption of macrolide antibiotics on erythromycin resistance in group A streptococci in Finland. Finnish Study Group for Antimicrobial Resistance. *N Engl J Med* 1997; **337**:441–446.
4. Gross R, Morgan AS, Kinky DE, Weiner M, Gibson GA, Fishman NO. Impact of a hospital-based antimicrobial management program on clinical and economic outcomes. *Clin Infect Dis* 2001; **33**:289–295.
5. Bamberger DM, Dahl SL. Impact of voluntary vs enforced compliance of third-generation cephalosporin use in a teaching hospital. *Arch Intern Med* 1992; **152**:554–557.
6. Fraser GL, Stogsdill P, Dickens Jr J-D, Wennberg DE, Smith Jr R-P, Prato BS. Antibiotic optimization. An evaluation of patient safety and economic outcomes. *Arch Intern Med* 1997; **157**:1689–1694.

7. Solomon DH, Van Houten L, Glynn RJ, *et al.* Academic detailing to improve use of broad-spectrum antibiotics at an academic medical center. *Arch Intern Med* 2001;161:1897–1902.
8. Gyssens IC. Quality measures of antimicrobial drug use. *Int J Antimicrob Agents* 2001;17:9–19.
9. Nathwani D, Tice A. Ambulatory antimicrobial use: the value of an outcomes registry. *J Antimicrob Chemother* 2002;49:149–154.
10. Gyssens IC, Blok WL, van den Broek PJ, Hekster YA, van der Meer JW. 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;16:904–912.
11. Gomez J, Conde-Cavero SJ, Hernandez-Cardona JL, *et al.* The influence of the opinion of an infectious disease consultant on the appropriateness of antibiotic treatment in a general hospital. *J Antimicrob Chemother* 1996;38:309–314.
12. Meyer KS, Urban C, Eagan JA, Berger BJ, Rahal JJ. Nosocomial outbreak of Klebsiella infection resistant to late-generation cephalosporins. *Ann Intern Med* 1993;119:353–358.
13. Dranitsaris G, Spizzirri D, Pitre M, McGeer A. A randomized trial to measure the optimal role of the pharmacist in promoting evidence-based antibiotic use in acute care hospitals. *J Technol Assess Health Care* 2001;17:171–180.
14. Marrie TJ, Lau CY, Wheeler SL, Wong CJ, Vandervoort MK, Feagan BG. A controlled trial of a critical pathway for treatment of community-acquired pneumonia. CAPITAL Study Investigators. Community-Acquired Pneumonia Intervention Trial Assessing Levofloxacin. *JAMA* 2000;283:749–755.
15. Clark LC, Davis CW. Experiences at a large teaching hospital with levofloxacin for the treatment of community-acquired pneumonia. *Am J Health Syst Pharm* 2000;57(Suppl. 3):S10–S13.
16. Jones SR, Pannell J, Barks J, *et al.* The effect of an educational program upon hospital antibiotic use. *Am J Med Sci* 1977;273:79–85.
17. Filius PMG, Bosboom RW, Groeneveld PHP, *et al.* Optimization of antibiotic use. Intensive monitoring in a collaboration (in Dutch). *Pharm Weekbl* 2000;135:254–258.
18. Cregin R, Segal-Maurer S, Weinbaum F, *et al.* Multidisciplinary approach to improving treatment of community-acquired pneumonia. *Am J Health Syst Pharm* 2002;59:364–368.
19. Natsch S, Steeghs MHM, Hekster YA, *et al.* Use of fluconazole in daily practice: still room for improvement. *J Antimicrob Chemother* 2000;48:303–310.
20. Fijn R, Chow M, Schuur PMH, de Jong-van den Berg LTW, Brouwers JRB. Multicentre evaluation of prescribing concurrence with anti-infective guidelines: epidemiological assessment of indicators. *Pharmacoepidemiol Drug Saf* 2002;11:361–372.
21. Taylor CT, Stewart LM, Byrd DC, Church CO. Reliability of an instrument for evaluating antimicrobial appropriateness in hospitalized patients. *Am J Health Syst Pharm* 2001;58:242–246.
22. Antibiotica werkgroep AZG. *Guidelines: antimicrobial therapy and prophylaxis* 1999. Groningen: AZG; 1999.
23. Kunin CM, Tupasi T, Craig WA. Use of antibiotics. A brief exposition of the problem and some tentative solutions. *Ann Intern Med* 1973;79:555–560.
24. Gyssens IC, van den Broek PJ, Kullberg BJ, Hekster Y, van der Meer JW. Optimizing antimicrobial therapy. A method for antimicrobial drug use evaluation. *J Antimicrob Chemother* 1992;30:724–727.
25. Altman DG. *Practical statistics for medical research*. 1st ed. London: Chapman & Hall; 1991.
26. Popping R. More agreeable software? *Comput Stat Data Anal Stat Software* 1992;13:223–227.
27. Schouten HJA. Measuring pairwise interobserver agreement when all subjects are judged by the same observers. *Stat Neerland* 1982;36:45–61.
28. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551–558.
29. Polderman KH, Jorna EMF, Girbes ARJ. Inter-observer variability in APACHE II scoring: effect of strict guidelines and training. *Intensive Care Med* 2001;27:1365–1369.
30. Gross ST. The kappa coefficient of agreement for multiple observers when the number of subjects is small. *Biometrics* 1986;42:883–893.
31. Carey JW. Intercoder agreement in analysis of responses to open-ended interview questions: examples from tuberculosis research. *Cultur Anthropol Methods* 1996;8:1–5.