ORIGINAL

Carolina A. M. Schurink
Christianne A. Van Nieuwenhoven
Jan A. Jacobs
Maja Rozenberg-Arska
Hans C. A. Joore
Erik Buskens
Andy I. M. Hoepelman
Marc J. M. Bonten

Clinical pulmonary infection score for ventilator-associated pneumonia: accuracy and inter-observer variability

C. A. M. Schurink (✉)
C. A. V. Nieuwenhoven · J. A. Jacobs
M. Rozenberg-Arska · H. C. A. Joore
E. Buskens · A. I. M. Hoepelman
M. J. M. Bonten
Department of Medicine,
Division of Acute Medicine
and Infectious Diseases,
Eijkman-Winkler Laboratory
for Microbiology,
Julius Centre for Health Sciences
and Primary Health Care,
University Hospital Maastricht,
University Medical Centre Utrecht
and Department of Medical Microbiology,
PO Box 85500,
HP F02.1263508 GA Utrecht,
The Netherlands
e-mail: K.Schurink@azu.nl
Tel.: +31-30-2506228
Fax: +31-30-2523741

Abstract *Objective:* Although quantitative microbiological cultures of samples obtained by bronchoscopy are considered the most specific tool for diagnosing ventilator-associated pneumonia, this labor-intensive invasive technique is not widely used. The Clinical Pulmonary Infection Score (CPIS), a diagnostic algorithm that relies on easily available clinical, radiographic, and microbiological criteria, could be an attractive alternative for diagnosing ventilator-associated pneumonia. Initially, the CPIS scoring system was validated upon 40 quantitative cultures of bronchoalveolar lavage fluid from 28 patients, and only few other studies have evaluated this scoring system since then. Therefore, little is known about the accuracy of this score. *Design:* We compared the scores of a slightly adjusted CPIS with results from quantitative cultures of bronchoalveolar lavage fluid in 99 consecutive patients with suspicion of ventilator-associated pneumonia, using growth of $\geq 10^4$ cfu/ml in bronchoalveolar lavage fluid as a cut-off for diagnosing ventilator-associated pneumonia. In addition, the CPIS were calculated for 52 patients by two different intensivists to determine the inter-observer variability. *Results:* Ventilator-associated pneumonia was diagnosed in 69 (69.6%) patients. When using a CPIS >5 as diagnostic cutoff, the sensitivity of the score was 83% and its specificity was 17%. The area under the Receiver Operating Characteristic curve was 0.55. The level of agreement for prospectively measured Clinical Pulmonary Infection Score (≤ 6 and >6) was poor (kappa =0.16). *Conclusions:* When compared to quantitative cultures of bronchoalveolar lavage fluid, the CPIS has a low sensitivity and specificity for diagnosing ventilator-associated pneumonia with considerable inter-observer variability.

Keywords Clinical pulmonary infection score · Ventilator-associated pneumonia · Intensive care unit · Inter-observer variability · Diagnosis

Introduction

Ventilator-associated pneumonia (VAP) is the most frequently occurring nosocomial infection in Intensive Care Unit (ICU) patients and has been associated with increased morbidity, prolonged duration of ventilation and ICU-stay, and increased costs for health care [1]. Diagnosing VAP is difficult and usually based on criteria with high sensitivity and low specificity [2] such as fever, leukocytosis, and infiltrative abnormalities on chest radiographs. Improvement of the diagnostic armamentarium for VAP has become an important research topic in intensive care research, and two different approaches have emerged [3, 4]; our capacity to establish a microbiological diagnosis, or improving our clinical judgment based on easily obtainable parameters.

The first approach includes invasive diagnostic techniques with quantitative microbiological analysis of samples obtained by bronchoalveolar lavage (BAL) or protected specimen brush (PSB). The second approach, also called the non-invasive strategy, relies on clinical criteria without microbiological confirmation by bronchoscopically obtained samples. When compared to a non-invasive approach, an invasive strategy based on direct examination of bronchoscopically samples for the presence of intracellular bacteria and quantitative cultures was associated with better survival and less antibiotic use [5]. However, bronchoscopy includes a certain, though probably small, risk for complications and processing specimens is expensive and labor intensive.

A reliable algorithm for clinical judgment based on easily available parameters would overcome these potential drawbacks. For example, the Clinical Pulmonary Infection Score (CPIS) described by Pugin and coworkers [6] consists of six easily obtainable clinical and laboratory variables. The CPIS was validated retrospectively in a population of 28 patients, with and without clinical suspicion of VAP, undergoing 40 BAL procedures. When compared to quantitative cultures of samples obtained by BAL, a CPIS score >6 had an excellent predictive value for diagnosing VAP. Since then, the CPIS has become a widely used diagnostic tool for VAP, both in clinical practice and in clinical trials [2, 6, 7, 8, 9, 10, 11, 12]. Optimally, such a widely used diagnostic algorithm should have a high diagnostic accuracy and low inter-observer variability, but such data are scarce or even completely absent. The aims of the present study, consisting of two substudies, were to compare CPIS scores with quantitative cultures of BAL fluid in patients with a clinical suspicion of VAP, using quantitative cultures of BAL as gold standard for VAP (study 1), and to determine inter-observer variability of this scoring system (study 2).

Methods

Setting

Study 1 was performed in three ICUs of two university hospitals in the Netherlands; the medical ICU (ten beds) and neurosurgical ICU (seven beds) of the University Medical Centre Utrecht (UMCU) and a mixed ICU of the University Hospital Maastricht (UHM) (15 beds). Data were collected from 1 June 1999 through 1 March 2001 in Utrecht and from 1 January 1997 through 31 December 1999 in Maastricht. Study 2, measuring the inter-observer variability, was only performed in Utrecht.

Data collection/study population

Study 1

All adult (age ≥18 years) patients mechanically ventilated for at least 48 h with a suspicion of VAP, according to the diagnostic cri-

teria for VAP (radiographic appearance of a new or progressive pulmonary infiltrate, and two of the three following: (1) fever, (2) leukokytosis, and (3) purulent tracheobronchial secretions [13]) and in which bronchoscopy could be performed on day of VAP suspicion, were eligible for the study. In case of concomitant antibiotic use, only patients who received antibiotics for >48 h and in whom the clinical suspicion of VAP had developed during antimicrobial treatment were included. The following data were routinely and prospectively recorded in computerized patient charts: body temperature, blood leukocyte count and number of band forms, character of tracheal secretions [purulent (yellow-green colour, with more than 25 leukocytes/ high power field) or not], microscopic examination (Gram stain) of tracheal secretions, semi-quantitative culture of bronchial secretions, ratio of arterial oxygen tension and inspiratory fraction of oxygen (PaO_2/FiO_2), interpretations of chest X-ray, and the use of antibiotics. When the variable was measured more than once in the 24 h, the value rendering the highest score (e.g., the most abnormal value) in 24 h prior to time of bronchoscopy was used. In addition, gender, age, and number of days on the ventilator prior to bronchoscopy were recorded.

Bronchoalveolar lavage and microbiology

BAL was performed by wedging the tip of the bronchoscope in the bronchus where the infiltrate was suspected, and infusing four 50 ml (UHM) or five 20 ml (UMCU) aliquots of sterile saline [14]. The first aliquot was discarded, and subsequent aliquots were pooled. In the UMH, quantitative cultures of BAL-fluid were performed by plating 2 μl and 10 μl on different culture media and in the UMCU by preparing 100-fold serial dilutions of BAL-fluid. Isolated bacteria were identified by standard laboratory methods and susceptibility to antimicrobial agents were determined according to National Committee for Clinical Laboratory Standards (NCCLS) methods[15]. Results were expressed in colony forming units per ml (cfu/ml).

Quantification of culture results was expressed as the log count and the bacterial index (BI). The log count was calculated by converting the sum of absolute counts of bacterial concentration(s) to the logarithmic number. The BI was calculated by summing the logarithmic numbers of individual log counts of bacterial concentrations. For example, a sample growing two bacterial species in concentrations of 1,000 cfu/ml each has a log count of 3.30 and a BI of 6.

Study 2

For measurement of inter-observer variability, two different intensivists calculated CPIS scores for 52 patients in the UMC Utrecht. In 42 patients, CPIS scores were calculated prospectively over a period of 3 months. On three randomly chosen days, CPIS of all patients in the ICUs were determined. A clinical suspicion of VAP was not required for these patients. The remaining ten patients were selected on the basis of a clinical suspicion of VAP in whom BAL had been performed (these patients were also included in study 1). In these patients, CPIS was calculated retrospectively by two intensivists. All intensivists had a list with CPIS definitions according to Pugin's definition (Table 1) when calculating scores.

Because of the disappointing results of study 2, the investigation was repeated with an adjusted flow-chart (Table 2). For this analysis CPIS were determined prospectively in 46 patients.

Table 1 Clinical Pulmonary Infection Score used for the diagnosis of ventilator-associated pneumonia [6]

Criterion	Range	Score
Temperature, °C	≥36.5 and ≤38.4	0
	38.5 and ≤38.9	1
	≥39 and ≤36	2
Blood leukocytes, mm³	≥4,000 and ≤11,000	0
	<4,000 or >11,000	1
	+band forms ≥500	2
Oxygenation, PaO_2/FiO_2	>240 or ARDS	0
mmHg	≤240 and no evidence of ARDS	2
Pulmonary radiography	No infiltrate	0
	Diffused (or patchy) infiltrate	1
	Localized infiltrate	2
Tracheal secretions	<14+[a] of tracheal secretions	0
	≥14+[a] of tracheal secretions	1
	+ purulent secretion	2
Culture of tracheal aspirate (semi-quantitative: 0–1–2 or 3+)	Pathogenic bacteria cultured ≤1+ or no growth	0
	Pathogenic bacteria cultured >1+	1
	+ same pathogenic bacteria seen on the Gram stain >1+	2

[a] Explained by the authors in the text as "quantity of tracheal aspirations: for each endotracheal aspiration the nurses estimated the quantity of secretions from 0 to 4+s; estimation of the volume of total secretions per day was calculated by adding all the + values recorded over 24 h together; to ensure a certain uniformity of this estimation only nurses having at least 2 years of continuous training in our ICU participated in this investigation"

Table 2 Adjusted and precise defined variables used in the second part of the inter-observer variability study. (*ARDS* adult respiratory distress syndrome)

CPIS points	0	1	2
Temperature, °C	≥36.1 and <38.4	>38.5 and <38.9	≤36.0 or ≥39.0
When temp measured orally or in groin: add 0.5°C.		In case of external cooling of the patient give 1 point	
Blood leukocytes, per 10⁹/l	≥4.0–≤11.0	≤3.9 ≥11.1 and in differentiation absence of band forms ≥11.1, no differentiation done and leukocytes ≤17.0	≥11.1 and in differentiation presence of band forms ≥11.1, no differentiation done and leukocytes ≥17.1
Tracheal secretions	Absence	Presence and non-purulent (=color: white or light-yellow)	Presence and purulent (=color: yellow, green or brown)
Oxygenation [PaO_2 (in mm Hg) x 100 / FiO_2 (in %)]	>240		<240 and no ARDS
Chest X-ray	No infiltrate	Diffused or patchy infiltrate	Localized infiltrate
Culture of tracheal aspirate (semi-quantitative: <10, 10–100, >100)[a]	<10	≥10 and ≤100	>100

General remarks: always use the most abnormal variable in the 24 h prior to time of scoring and in case of a missing variable, use the variable from the previous day. When this variable is also missing give 0 points

[a] Take previous tracheal aspirate culture or sputum during this admission; when no previous culture done: 0 points

Definitions

Clinical pulmonary infection score (CPIS)

The CPIS as described by Pugin et al. included six variables: (1) body temperature, (2) blood leukocyte count and number of band forms, (3) character of tracheal secretions (purulent or not) and quantity of tracheal aspirates, (4) microscopic examination (Gram stain) and semi- quantitative culture results of the bronchial secretions, (5) arterial oxygen tension/inspiratory fraction of oxygen (PaO2/FiO2), and (6) chest X-ray [6] (Table 2). In the original study, quantities of tracheal aspirates were measured by trained ICU-nurses who scored quantities of secretions from 0 to 4+. The total volume of secretions per day was calculated by adding all the + values. Since the nurses in our ICUs were not specifically trained to score quantities of tracheal aspirates and amounts were not uniformly registered, this variable was not included in study 1. Therefore, the possible scores for that part of our study ranged from 0 to 10.

In the analysis of the inter-observer variability (study 2) the variable quantities of tracheal aspirates were, in fact, included. The estimated quantities of tracheal aspirates noted by nurses were

Table 3 Patient characteristics (*n*=99)

	Non-microbiologically proven VAP[a] *n*=30	Microbiologically proven VAP[a] *n*=69
Male sex, *n* (%)	16 (53)	52 (75)
Age (years)[b]	61±15 (30–85)	63±15 (28–87)
Duration of ventilation until VAP (days)[b]	12.3±13.8 (3–65) median 6	12.4±10.1 (3–43) median 9
No of patients receiving antibiotics in 24 h prior to BAL, *n* (%)	7 (23)	26 (38)
In ICU mortality, *n* (%)	12 (40)	24 (35)
Admitting condition:		
Cardiac failure, *n* (%)	2 (7)	3 (4)
Neurologic disease, *n* (%)	3 (7)	10 (14)
Trauma, *n* (%)	2 (7)	9 (13)
Post cardiac surgery, *n* (%)	2 (7)	5 (7)
Other post-surgery, *n* (%)	5 (17)	9 (13)
Pneumonia, *n* (%)[c]	4 (13)	21 (30)
Miscellaneous, *n* (%)	7 (40)	12 (17)

[a] Defined as growth quantative culture BAL ≥10⁴ cfu/ml
[b] Values are means ± SD and (range),
[c] Either community or hospital acquired pneumonia, cured before development of VAP

interpreted by the intensivists and translated in 0–4+ as described by Pugin in the first part of this study.

Statistical analysis

Data were expressed as absolute numbers with percentages and as means or medians with standard deviation or ranges. Comparisons were performed by chi-square or t-test, when appropriate. A probability value less than 0.05 was considered statistically significant.

Using the results of quantitative cultures of BAL (BI ≥5 and log count ≥4) as "gold standard", Receiver Operator Characteristic (ROC) curves correlating true and false positive rates (sensitivity and 1-specificity, respectively) for the different "CPIS" thresholds were constructed. Univariate and multivariate logistic regression analysis were used to determine the association between the individual scoring variables and VAP (defined as log count ≥4). The variables were dichotomized into 'normal' (CPIS variable 0) and 'abnormal' (CPIS variable 1 and 2). The inter-observer variability was analyzed for the individual variables of the CPIS as well as for total scores (using a cut-off point of >6) by calculating the kappa coefficient, a dimension of clinical agreement of two observers [16].

Results

Study 1

Patient characteristics

A total of 99 patients (68.7% male) were included (89 in Maastricht and ten in Utrecht), with a mean age of 62±15 years (range 28–87 years) (Table 3). The mean

Fig. 1a,b **a** Scatter plot of comparison between log count of BAL fluid and Clinical Pulmonary Infection Score; **b** Scatter plot of comparison between Bacterial Index of BAL fluid and Clinical Pulmonary Infection Score
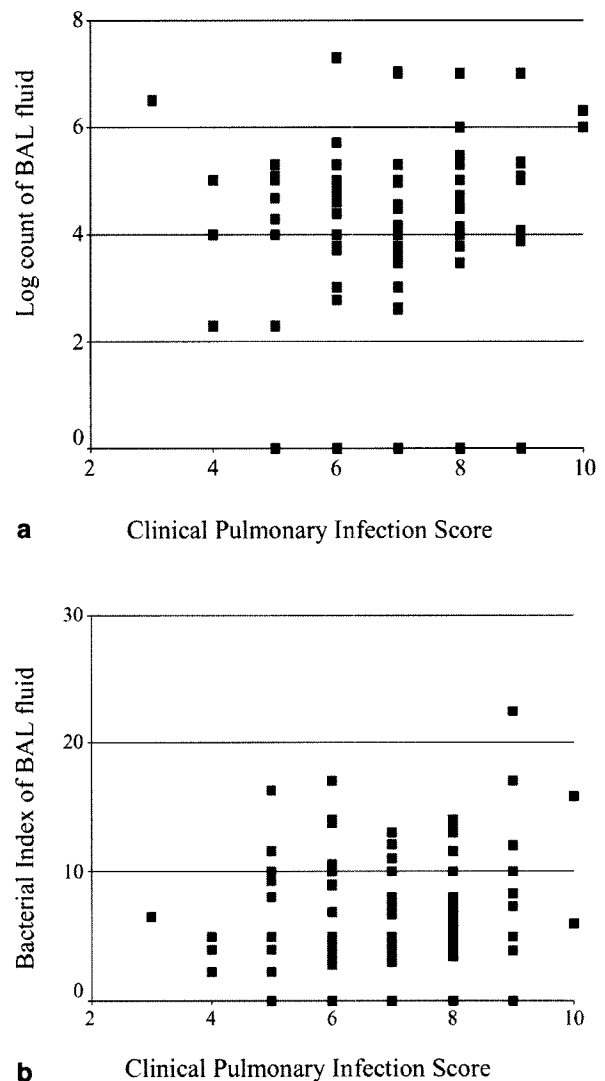

a Clinical Pulmonary Infection Score

b Clinical Pulmonary Infection Score

Fig. 2 ROC of Clinical Pulmonary Infection Score and quantitative culture of BAL fluid with Bacterial Index ≥5 and log count ≥4
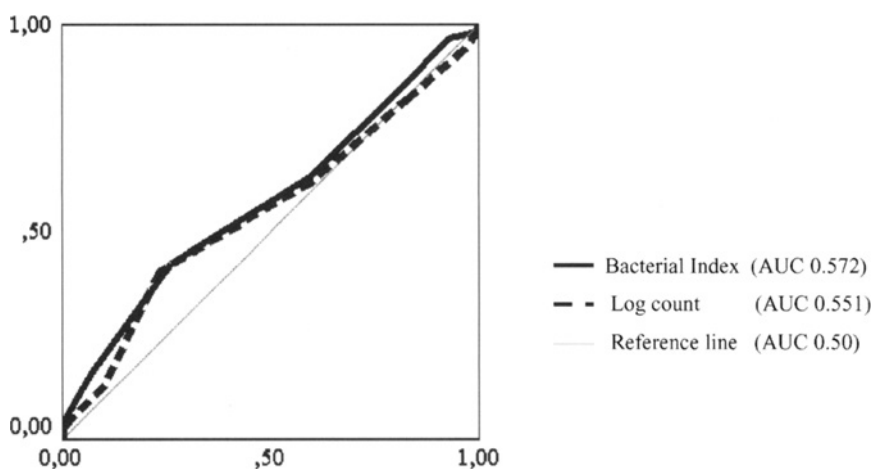

Legend:
— Bacterial Index (AUC 0.572)
− − Log count (AUC 0.551)
Reference line (AUC 0.50)

Table 4 Association between individual variables of Clinical Pulmonary Infection Score and Ventilator-Associated Pneumonia. (VAP defined as ≥10^4 cfu/ml in BAL)

CPIS variable[a]	Univariate				Multivariate			
	Crude Odds	95% CI		P-value	Adjusted Odds	95% CI		P-value
		Lower	Upper			Lower	Upper	
Temperature	1.092	0.431	2.770	0.853	1.032	0.384	2.776	0.950
Blood leukocytes	2.340	0.790	6.929	0.125	1.945	0.619	6.107	0.255
Culture	4.444	1.615	12.229	0.004	4.249	1.474	12.245	0.007
Oxygenation	1.655	0.496	5.522	0.413	2.214	0.602	8.143	0.232

[a] The variables were dichotomized into normal (0) and abnormal (1 or 2).
The variable Chest X-ray, is by definition according to the CDC criteria, in all cases abnormal, and is not included in this analysis

duration of mechanical ventilation at the time of bronchoscopy was 12.4±11.3 days (range 3–65 days). Based on quantitative cultures of BAL, the clinical suspicion of VAP was microbiologically confirmed in 69 patients (69.7%). Thirty-three (33.3%) patients received antibiotics at the time of bronchoscopy, all during >48 h and in all patients a new clinical suspicion of VAP developed while receiving antibiotics. The clinical suspicion was microbiologically confirmed in 26 of these 33 patients. Thirty-six (36.4%) patients died in the ICU.

Comparison of CPIS and quantitative cultures of BAL

Because the variable "quantities of secretions" was not available, the highest possible score of CPIS was 10. CPIS scores varied from 3 to 10 (median 7; mean 6.9±1.4). BI values of BAL samples varied from 0 to 22.4 (mean 6.4±4.6) and log counts varied from 0 to 7.3 (mean 4±1.9). Correlations between CPIS values and BI or log count values appeared to be poor (r=0.178, P=0.079 and r=0.115, P=0.257 for CPIS and BI and CPIS and log count, respectively) (Fig. 1a and Fig. 1b).

When using a log count ≥4 of BAL samples as gold standard, CPIS value >7, had the largest area under the curve (AUC) in the Receiver Operator Characteristic (ROC) curve (0.644). The AUC for ROC curves with CPIS>6 and CPIS>8 and log count≥4 were 0.541 and 0.640, respectively. Sensitivity and specificity of CPIS>7 were 41% and 77%, respectively, with a positive predictive value of 80% and a negative predictive value of 36%. Comparable results were obtained for ROC analysis of CPIS scores and BI values (Fig. 2). Exclusion of patients admitted with pneumonia or those receiving antibiotics at the time of bronchoscopy did not change the results (data not presented).

Of the individual CPIS variables, only a positive microbiological culture of tracheal aspirate was significantly associated with VAP (defined as log count ≥4) (Odds 4.25, 95% CI 1.474–12.25; P=0.007) (Table 4). The variable chest X-ray was not evaluated in logistic regression analysis because this variable is abnormal in all cases, by definition of CDC criteria.
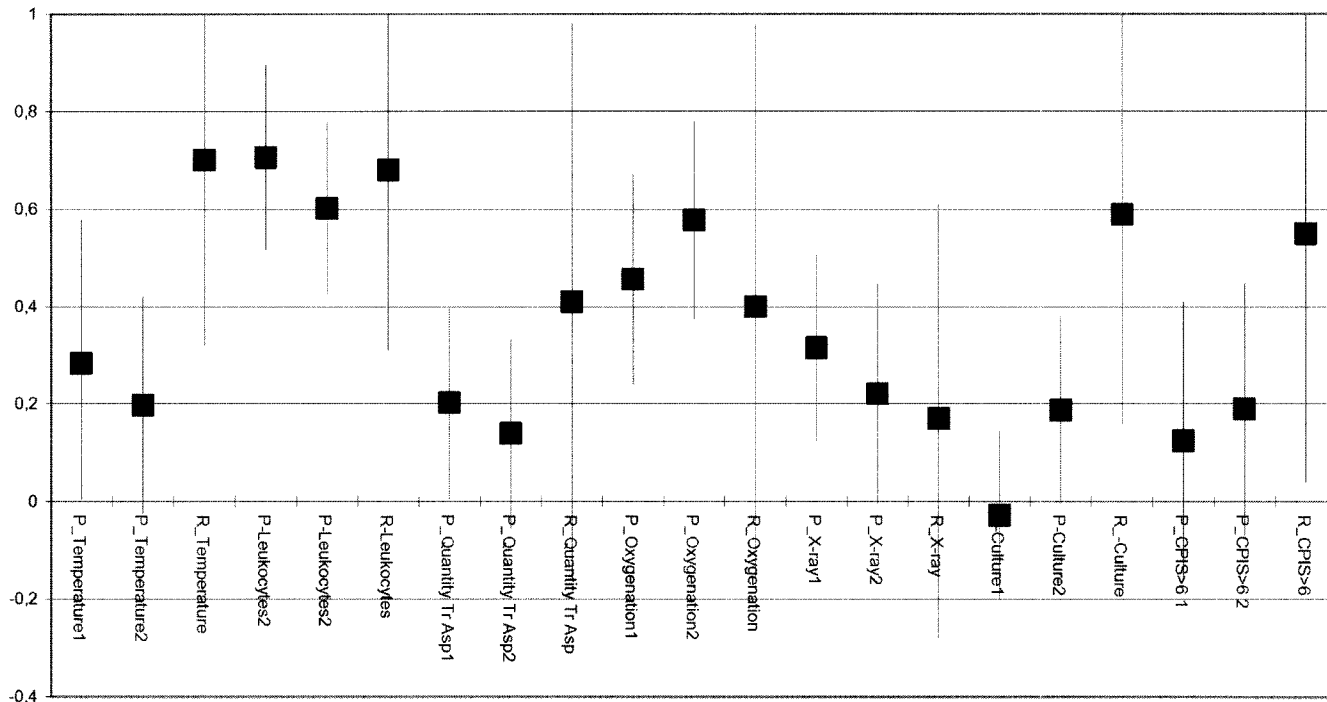
Fig. 3 Inter-observer variability (kappa) for the different Clinical Pulmonary Infection Score-variables and for CPIS >6 measured prospectively (P_1) (n=42), prospectively with more defined criteria (P_2) (n=46), retrospectively (R) (n=10) with 95% CI

Study 2

Agreement between observers in measuring individual CPIS variables varied from 0.02 for culture results in prospective analysis to 0.7 for temperature in retrospective analysis (Fig. 3). Importantly, physicians frequently decided that variables could not be scored as data were not available at that time point. This could be due to pending data of culture results, impossibility to score body temperature between 36.0 and 36.5, as this was not defined in the original CPIS, or due to missing data, e.g., when chest X-ray was not performed on that day. Although some physicians used the variables of the previous day in case of missing values or gave an interpretation of temperature between 36 and 36.5, others found that they could not use these data to score CPIS.

The CPIS scores (maximum score of 12) of 42 randomly chosen patients with or without a clinical suspicion of VAP varied from 0 to 8 (mean 3.26±1.91), with a level of agreement (kappa) for CPIS <6 and ≥6 of 0.16. This extremely low level of agreement is explained by the inclusion of missing data for individual scoring variables. When excluding patients with missing variables kappa for CPIS>6 and ≤6 was 0.6 (data not shown). The retrospective calculation of CPIS scores in patients with

a clinical suspicion of VAP (that had undergone bronchoscopy) varied from 3–10 (mean 7.9±2.46) with a level of agreement for CPIS>6 and ≤6 of 0.55.

In attempt to decrease inter-observer variability, a flowchart (Table 2) was developed. Now kappa values varied from 0.13 for quantity of tracheal secretions to 0.6 for leukocyte counts, with a level of agreement (kappa) for CPIS ≤6 and >6 of 0.18 (Fig. 3). When excluding patients with missing variables, kappa was 0.5 for CPIS>6 and ≤6.

Discussion

Associations between CPIS and quantitative cultures of BAL samples in patients with a clinical suspicion of VAP appeared to be poor and the level of agreement between different physicians for scoring individual variables varied greatly. These findings question the reliability of the CPIS as a diagnostic tool for VAP.

The absence of a true gold standard for VAP hampers the interpretation of any study investigating the diagnostic approach of this infection, and is also a confounder of the present study. We have chosen to use quantitative cultures of BAL as surrogate gold standard, fully realizing that sensitivity and specificity of BAL cultures are not 100%. Moreover, since Pugin et al. validated the CPIS scores on bacterial index of log counts, we have used both the log count of microorganisms per ml as well as the bacterial index for comparison with CPIS score. The bacterial index was introduced by Johanson et al. [17], but is probably not widely used nowadays.

Validation of CPIS has been attempted in seven studies, but due to differences in comparative diagnostic techniques that were used and modifications of the original CPIS criteria, direct comparison of these studies is difficult. In their original analysis Pugin and coworkers [6] found an excellent correlation between CPIS>6 and bacterial index >5 of quantitative cultures obtained by BAL. In their analysis of 40 BAL procedures in 28 patients, a CPIS >6 had a sensitivity of 93% and a specificity of 96% for diagnosing VAP, which corresponds to an area under the ROC curve of 0.95. In two other studies the original CPIS was compared to post mortem results with VAP diagnosed upon histological criteria [2, 8]. Using a CPIS>6 as cutoff point, sensitivities were 72% and 77%, and specificities 85% and 42%, respectively. The comparison of the unmodified CPIS to the consensus of two investigators in a study of 59 children with suspected VAP gave an area under the ROC curve of 0.81 [9].

In the remaining three studies, the original CPIS was modified. A'Court et al. [7, 11] deleted the variable 'culture result' and added three criteria: [1] clinical course on/off antibiotics consistent with pneumonia, [2] lack of evidence for an alternative source of sepsis, and [3] lung biopsy or post-mortem histology demonstrating pneumonia within a relevant time span. VAP was confirmed retrospectively either by CPIS>8 plus one out of three additional criteria, or by CPIS>6 plus two out of three additional criteria. Compared to quantitative cultures of BAL (cut off $\geq 10^4$ cfu/ml) the sensitivity of this modified CPIS was 93%. Flanagan and colleagues [10] deleted the variable 'culture result' and used a different description of sputum quantity, which was divided into scanty (0 points), moderate/profuse but non-purulent (1 point) or moderate/profuse and purulent (2 points). The adjusted CPIS was compared to a clinical suspicion of VAP according to CDC criteria, which was considered confirmed by histopathological confirmation or by concurrent isolation of a pathogen from distal respiratory samples, pleural fluid culture or blood culture. In an analysis of 34 episodes of clinically suspicions of VAP in 32 patients, CPIS>7 as a cutoff point had a sensitivity of 85% and specificity of 91%. In the most recent study, a modified CPIS (using five variables at day 1: temperature, leukocytes, secretions, PaO_2/FiO_2, chest X-ray) had a sensitivity of 60% and specificity of 59% for diagnosing VAP when using a value >6 as cutoff. Sensitivity and specificity improved somewhat (to 78% and 56%, respectively), when Gram stain results of BAL fluid were added to the score [18].

How can we explain the poor correlations found in the present study, as compared to better associations in some other studies? Patient selection is a possibility. The pre-test probability of VAP may have been higher in our population as only adult patients with VAP according to

CDC criteria were included. In other studies patients without a clinical suspicion [6], or children [9] have been included. Recently, Michaud and coworkers demonstrated in a meta-analysis of studies on diagnostic tests for VAP that patient selection (i.e., inclusion of patient with a clinical suspicion of VAP) had the largest effect on the measurement of a test's performance [19]. Inclusion of patients without a clinical suspicion of VAP could lead to false-positive results. They also recommended to use BAL volumes ≥ 140 ml and to obtain pulmonary secretions before start of antibiotics to get a higher performance of diagnostic tests. In ten patients included in our study (those studied in Utrecht) BAL was performed with <140 ml lavage fluid, which might have negatively influenced our results. Another explanation could be the "gold standard" that was used. In two studies, CPIS scores were not calculated at the time of infection, but at the time of death with post-mortem examination as a gold standard [2, 8]. Finally another explanation could be the sample size, as smaller patient populations, ranging from 25 to 59 patients were included in all other studies [2, 6, 7, 8, 9, 10].

As compared to the original CPIS score defined by Pugin, we had to introduce some adjustments. Since quantities of tracheal aspirates were not monitored routinely by our nursing staff this variable was omitted, precluding an optimal comparison of CPIS test characteristics to those reported by Pugin. This being said, low inter-observer variability also is a prerequisite for general use of a diagnostic test. In our study the agreement of individual CPIS variables between observers was low, especially for frequently measured variables, and the introduction of a flowchart did hardly improve agreement levels. However, the high number of missing data remains an insufficiency of our study.

Positive microbiological cultures of tracheal aspirates obtained at the day of the clinical suspicion of VAP, appeared to be the best predictor of VAP. However, these results would only be available 24–48 h later and this variable is, therefore, hardly suitable for prospective use. Although surveillance cultures of tracheal aspirates are frequently used for selection of empirical treatment, its benefits on accuracy of empirical therapy in case of VAP have not been determined. The positive predictive value of surveillance cultures to identify the causative microorganism causing VAP was only 18% in a recent study [20].

The CPIS has been successfully used to guide duration of therapy in some patients and as a tool to investigate the effects of infection prevention. Singh et al. [12] used an adjusted CPIS score (counting five variables: temperature, leukocytes, secretions, PaO_2/FiO_2, chest X-ray) to safely discontinue antimicrobial therapy after 3 days in patients with persistent low CPIS scores (≤ 6). In addition, CPIS may be used for longitudinal analysis of patients in intervention studies to prevent VAP [21].

However, our findings on the inter-observer variability justify that as few as possible persons calculate CPIS scores.

In conclusion, when compared to quantitative cultures obtained by BAL, specificity and sensitivity of CPIS appeared to be low. Moreover, inter-observer variability for individual CPIS variables and overall score of CPIS was low. These findings justify a cautious use of CPIS as a diagnostic tool for VAP.

References

1. Kollef MH (1999) The prevention of ventilator-associated pneumonia. N Engl J Med 340:627–634
2. Papazian L, Thomas P, Garbe L, Guignon I, Thirion X, Charrel J, Bollet C, Fuentes P, Gouin F (1995) Bronchoscopic or blind sampling techniques for the diagnosis of ventilator-associated pneumonia. Am J Respir Crit Care Med 152:1982–1991
3. Niederman MS, Torres A, Summer W (1994) Invasive diagnostic testing is not needed routinely to manage suspected ventilator-associated pneumonia. Am J Respir Crit Care Med 150:565–569
4. Chastre J, Fagon JY (1994) Invasive diagnostic testing should be routinely used to manage ventilated patients with suspected pneumonia. Am J Respir Crit Care Med 150:570–574
5. Fagon JY, Chastre J, Wolff M, Gervais C, Parer-Aubas S, Stephan F, Similowski T, Mercat A, Diehl JL, Sollet JP, Tenaillon A (2000) Invasive and noninvasive strategies for management of suspected ventilator-associated pneumonia. A randomized trial. Ann Intern Med 132:621–630
6. Pugin J, Auckenthaler R, Mili N, Janssens JP, Lew PD, Suter PM (1991) Diagnosis of ventilator-associated pneumonia by bacteriologic analysis of bronchoscopic and nonbronchoscopic "blind" bronchoalveolar lavage fluid. Am Rev Respir Dis 143:1121–1129
7. A'Court CH, Garrard CS, Crook D, Bowler I, Conlon C, Peto T, Anderson E (1993) Microbiological lung surveillance in mechanically ventilated patients, using non-directed bronchial lavage and quantitative culture. Q J Med 86:635–648
8. Fabregas N, Ewig S, Torres A, El-Ebiary M, Ramirez J, de La Bellacasa JP, Bauer T, Cabello H (1999) Clinical diagnosis of ventilator-associated pneumonia revisited: comparative validation using immediate post-mortem lung biopsies [see comments]. Thorax 54:867–873
9. Fischer JE, Janousek M, Nadal D, Fanconi S (1998) Diagnostic techniques for ventilator-associated pneumonia [letter; comment]. Lancet 352:1066–1067
10. Flanagan PG, Findlay GP, Magee JT, Ionescu A, Barnes RA, Smithies M (2000) The diagnosis of ventilator-associated pneumonia using non- bronchoscopic, non-directed lung lavages. Intensive Care Med 26:20–30
11. Garrard CS, A'Court CD (1995) The diagnosis of pneumonia in the critically ill. Chest 108:17S-25S
12. Singh N, Rogers P, Atwood CW, Wagener MM, Yu VL (2000) Short-course empiric antibiotic therapy for patients with pulmonary infiltrates in the intensive care unit. A proposed solution for indiscriminate antibiotic prescription. Am J Respir Crit Care Med 162:505–511
13. Pingleton SK, Fagon JY, Leeper KV Jr (1992) Patient selection for clinical investigation of ventilator-associated pneumonia. Criteria for evaluating diagnostic techniques. Chest 102:553S-556S
14. Chastre J, Fagon JY, Soler P, Bornet M, Domart Y, Trouillet JL, Gibert C, Hance AJ (1988) Diagnosis of nosocomial bacterial pneumonia in intubated patients undergoing ventilation: comparison of the usefulness of bronchoalveolar lavage and the protected specimen brush [published erratum appears in Am J Med 86:258]. Am J Med 85:499–506
15. NCCLS (2001) National Commitee for Clinical Laboratory Standards: Performance standards for antimicrobial susceptibility testing. Eleventh informational supplement, Wayne, PA NCCLS Document 100-S111
16. Cohen J (1960) A coefficient of agreement for nominal scales. Education Psych Measure 20:37–46
17. Johanson WG Jr, Seidenfeld JJ, Gomez P, de los Santos R, Coalson JJ (1988) Bacteriologic diagnosis of nosocomial pneumonia following prolonged mechanical ventilation. Am Rev Respir Dis 137:259–264
18. Fartoukh M, Maitre B, Honore S, Cerf C, Zahar JR, Brun-Buisson C (2003) Diagnosing pneumonia during mechanical ventilation: the Clinical Pulmonary Infection Score revisited. Am J Respir Crit Care Med 168:173–179
19. Michaud S, Suzuki S, Harbarth S (2002) Effect of design-related bias in studies of diagnostic tests for ventilator-associated pneumonia. Am J Respir Crit Care Med 166:1320–1325
20. Hayon J, Figliolini C, Combes A, Trouillet JL, Kassis N, Dombret MC, Gibert C, Chastre J (2002) Role of serial routine microbiologic culture results in the initial management of ventilator-associated pneumonia. Am J Respir Crit Care Med 165:41–46
21. Pugin J, Auckenthaler R, Lew DP, Suter PM (1991) Oropharyngeal decontamination decreases incidence of ventilator- associated pneumonia. A randomized, placebo-controlled, double-blind clinical trial. JAMA 265:2704–2710