JOURNAL OF CLINICAL MICROBIOLOGY, Oct. 1982, p. 627–631
0095-1137/82/100627-05$02.00/0

Vol. 16, No. 4

Comparison of Six Different Criteria for Judging the Acceptability of Sputum Specimens

LESTER K. WONG,* ARTHUR L. BARRY, AND SUSAN M. HORGAN

Department of Pathology, University of California, Davis, Medical Center, Sacramento, California 95817

Received 29 April 1982/Accepted 28 June 1982

A series of 391 unselected expectorated sputum specimens was examined microscopically, and six different published criteria for judging the acceptability of the specimens were applied. Of the 391 specimens, 234 were found to be acceptable or unacceptable by all six criteria; 157 specimens were discrepant. By the criteria of Murray and Washington and of Barry, 25 and 23% of the specimens, respectively, were rejected; only 19 of 143 specimens which contained potential pathogens as part of their predominant microbial flora were rejected by both sets of criteria. The criteria described by Geckler et al. and Bartlett missed fewer potential pathogens; only 9 or 17% of the specimens, respectively, were found unacceptable. The criteria of Heineman and Radano and of Van Scoy resulted in the greatest percentages of specimens judged unacceptable (28 and 29%, respectively), including 19 and 24% of specimens containing potential pathogens. The reproducibility of sputum screening results was also assessed, comparing the method of Murray and Washington with that of Barry. Six separate slides were prepared from each of 45 different specimens: three samples with purulent or bloody flecks and three samples in which the specimens had been mixed with an applicator swab. Satisfactory reproducibility was observed with both criteria and both sampling methods; no significant differences in reproducibility could be documented in this limited series.

Even under the best conditions, the results of bacteriological examination of expectorated sputum specimens are difficult to interpret. When potential pathogens are recovered, the physician must decide whether the isolate represents a true pathogen or one that is simply colonizing the oropharynx. That decision is often impossible since all expectorated sputum is contaminated with oropharyngeal flora, which may include potential pathogens. The amount of oropharyngeal contamination can be judged by examining the cellular components in a stained smear of the specimen. Since squamous epithelial cells are found only in the upper respiratory tract, their presence suggests oropharyngeal contamination. On the other hand, inflammatory cells (primarily polymorphonuclear leukocytes) suggest material derived from the site of an active infection.

In 1974, Bartlett (3) first suggested that clinical laboratories should examine sputum specimens microscopically and refuse to culture specimens showing evidence of excessive oropharyngeal contamination. He proposed a scheme for grading the quality of sputum specimens after a slide of the specimen had been examined at 100× magnification. His Q score was based on the relative number of squamous epithelial cells, inflammatory cells, and mucus seen in gram-stained smears. Specimens rejected by this scheme are those that are not likely to produce interpretable results because of excessive oropharyngeal contamination. If such specimens contain potential pathogens, it is difficult to determine whether the isolates are true pathogens or simply colonizing the upper airway.

In 1975, Murray and Washington (7) described a simpler scheme for judging the quality of sputum specimens. They rejected all specimens with more than 10 squamous epithelial cells per average low-power field (LPF). Van Scoy (8) reexamined their data and concluded that specimens with more than 25 leukocytes per LPF should be accepted regardless of the number of epithelial cells present. On the other hand, Geckler et al. (4) concluded that only specimens with $\leq$25 epithelial cells per LPF should be accepted regardless of the number of leukocytes. Because the cells are often unevenly distributed over the slide, inconsistencies might be anticipated unless the entire slide is examined, especially when one type of cell is being assessed. Also, inconsistencies might be anticipated if slides are prepared by sampling different purulent or bloody portions of the same specimen as recommended by Bartlett et al. (2). It is

TABLE 1. Summary of six published criteria for judging acceptability of sputum specimens

Author (reference)	Method[a]	Criteria for acceptability
Bartlett (3)	Assign + and − values; +2 if >25 WBC; +1 if 10–25 WBC; +1 if mucus seen; −2 if >25 EPI; −1 if 10–25 EPI	Any positive score (sum of + and − values assigned)
Murray and Washington (7)	Avg no. of EPI/LPF	<10 EPI/LPF
Geckler et al. (4)	Avg no. of EPI/LPF	<25 EPI/LPF
Van Scoy (8)	Avg no. of WBC/LPF	>25 WBC/LPF
Barry (1)	Assign + and − values: +3 if >150 WBC; +2 if 76–150 WBC; +1 if 1–75 WBC; −3 if >25 EPI; −2 if 16–25 EPI; −1 if 5–15 EPI	Any positive score (sum of + and − values assigned)
Heineman and Radano (6)	Avg ratio of WBC to EPI	>10 WBC/EPI

[a] Stained smears were examined under 100× magnification, and the number of squamous epithelial cells (EPI) or white blood cells (WBC) per LPF was determined.

difficult to be certain that the portion used for microscopic examination is the same as that used for culturing when only selected portions are examined.

In 1978, Barry (1) described another method for judging the quality of expectorated sputum. This scheme, like that of Bartlett (3), evaluates the ratio between white blood cells and squamous epithelial cells and therefore minimizes the effect of variations in the thickness of material in different areas of the same slide. In 1979, Heineman and Radano (6) described a similar scheme for screening sputum specimens: they rejected specimens with more than 1 epithelial cell for every 10 white blood cells in the average low power field. This represented a change from an earlier communication (5), which recommended that specimens with more than 1 epithelial cell for every 20 white blood cells be rejected.

In the present report, we directly compare the six different published criteria for judging the quality of sputum specimens and evaluate all six schemes against bacteriological examination of 391 consecutive sputum specimens submitted to our laboratory for examination. An additional 45 specimens were examined to compare the reproducibility of the criteria of Murray and Washington (7) and of Barry (1) and to compare the reproducibility of two sampling methods.

MATERIALS AND METHODS

The 391 unselected expectorated sputum specimens were examined by two of us (L.W. and S.H.). Each specimen was first mixed with an applicator swab and then inoculated onto blood agar, Fildes peptic digest agar, and MacConkey agar plates, and a smear was prepared for Gram staining. Each stained smear was examined microscopically under low-power (100×) magnification, and the cellular components were evaluated. Each specimen was then categorized as accepted or rejected by each of the six different criteria (Table 1). All of the inoculated media were processed regardless of the appearance of the stained smears. The amount of growth was evaluated by criteria defined elsewhere (1), and the predominant microorganisms were identified. Predominant or copredominant *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Haemophilus influenzae*, and other gram-negative bacilli were considered potential pathogens which might represent etiological agents or unusual oropharyngeal colonization.

Reproducibility studies were performed with 45 additional specimens selected for their nonhomogeneity. For each specimen six slides were prepared, three by selecting purulent or bloody flecks and three by sampling separate areas after the specimen was mixed with an applicator swab. All six samples from the first 20 specimens were examined bacteriologically; only one mixed sample from each of the remaining 25 specimens was cultured. Stained slides were coded and then examined by different technologists in a single-blind protocol, and the criterion of Barry (1) was applied: all of the technologists were accustomed to that system. All specimens with 5 to 15 epithelial cells per LPF were later reexamined for the criterion of Murray and Washington (7). The other criteria were not evaluated in this reproducibility phase of the study.

RESULTS

The evaluations of 391 specimens by the six criteria are summarized in Table 2. All six screening criteria provided the same interpretation for 234 of the 391 expectorated specimens (217 accepted and 17 rejected by all six criteria). The overall rejection rate was highest when the criteria of Heineman and Radano (28%) and Van Scoy (29%) were used and lowest when the criteria of Geckler et al. (9%) and Bartlett (17%) were used. The difference in the number of specimens rejected by these four criteria was statistically significant ($P < 0.01$, chi-square test). When the results obtained by using the criterion of Geckler et al. were excluded, no significant differences could be documented among the results obtained with the other three criteria.

Potential pathogens were recovered from 143 specimens. The percentage of specimens found

TABLE 2. Microscopic examination of 391 expectorated sputum specimens, applying six different criteria for acceptance or rejection

Method (criterion for acceptance)	Microscopic examination		No. (%)[a] of specimens with potential pathogens[b]:		Avg no. of species isolated (accepted/rejected)
	No. accepted	No. rejected (%)	Accepted	Rejected	
Geckler et al. (<25 EPI/LPF)	356	35 (9)	134 (38)	9 (27)	2.5/3.3
Bartlett (positive score)	326	65 (17)	127 (39)	16 (25)	2.5/3.0
Barry (positive score)	302	89 (23)	124 (41)	19 (21)	2.5/3.2
Murray et al. (<10 EPI/LPF)	292	99 (25)	124 (42)	19 (19)	2.4/3.2
Heineman et al. (>10 WBC/EPI)	280	111 (28)	116 (41)	27 (24)	2.4/3.2
Van Scoy (>25 WBC/LPF)	277	114 (29)	109 (39)	34 (30)	2.6/2.6
Agreement[b]	217	17 (11)	94 (43)	1 (6)	2.4/3.3

[a] The percentage of the total number accepted or rejected by each screening criterion.
[b] One or more interpretive discrepancies were observed with 157 specimens (23%); 48 (30%) of those contained potential pathogens, and an average of 2.8 species were isolated.

acceptable which contained potential pathogens varied from 38 (Geckler et al.) to 42% (Murray and Washington). Only one specimen with potential pathogens was rejected by all six criteria. A total of 27 specimens with potential pathogens were rejected when the criterion of Heineman and Radano (6) was used. By the criterion of Van Scoy, 34 specimens with potential pathogens were found unacceptable. The criteria of Murray and Washington (7) and Barry (1) both resulted in 19 specimens with potential pathogens being rejected (13% of specimens with potential pathogens, or 19 to 21% of all rejected specimens). Only 11% of the specimens with potential pathogens were rejected when the criterion of Bartlett (3) was used, and 6% were rejected when the criterion of Geckler et al. (4) was used; 25 to 27% of the few rejected specimens contained potential pathogens. The average number of species recovered was about the same for all six screening criteria: accepted specimens averaged about 2.5 predominating species, whereas rejected specimens contained at least 3 different predominating species. No difference in the number of predominating species between accepted and rejected specimens was found when the criterion of Van Scoy was used, i.e., 2.6 species were recovered from both types of specimens.

The screening criteria differed in recommending rejection or acceptance of 157 of the 391 specimens. Forty-eight (30%) of the discrepant specimens contained potential pathogens. The pattern of acceptance and rejection of the 157 specimens with discrepancies between screening criteria is shown in Table 3. The results obtained by using the criterion of Murray and Washington (7) generally agreed with those obtained by using the criterion of Barry (1): with the former, 27 specimens were rejected that were found acceptable by the latter, and 17 specimens were found acceptable by the former that were rejected by the latter. Bartlett's criterion excluded 48 specimens that were found acceptable by at least one of the other criteria. On the other hand, with the criteria of Heineman and Radano and of Van Scoy, respectively, 94 and 97 specimens were rejected that were acceptable by at least one of the other criteria.

The reproducibility of two sampling methods

TABLE 3. Pattern of acceptance or rejection of sputum specimens when six different published criteria were applied: summary of data with 157 specimens providing interpretive discrepancies

Criterion	No. of specimens accepted	No. of accepted specimens subsequently rejected by criterion of:					
		Geckler et al.	Bartlett	Barry	Murray et al.	Heineman et al.	Van Scoy
Geckler et al.	139		37	60	69	78	95
Bartlett	109	4		35	51	48	63
Barry	85	2	10		27	21	57
Murray et al.	75	1	15	17		28	69
Heineman et al.	63	1	4	2	19		46
Van Scoy	60	14	14	33	35	41	
No. of specimens subsequently rejected		18	48	72	82	94	97

TABLE 4. Reproducibility of two sampling methods and two interpretive criteria for judging the acceptability of sputum specimens

Sampling method[a]	Criterion	No. of specimens (n = 45):		
		Accepted[b]	Rejected[b]	Discrepant result
Selected flecks	Murray et al.	31	8	6
	Barry	28	8	9
Mixed sample	Murray et al.	30	13	2
	Barry	28	13	4

[a] All slides were prepared in triplicate.
[b] All three slides gave same results.

and two of the interpretive criteria were also evaluated (Table 4). The majority of specimens selected for this phase of the study were judged to be acceptable under both criteria. Only 8 specimens were consistently found unacceptable when the slides were prepared by selecting purulent or bloody flecks and 13 specimens were consistently unacceptable when the specimens were first mixed. Reproducibility can be assessed by comparing the number of specimens showing discrepancies among triplicate slides. The results from slides prepared from mixed specimens tended to be the most reproducible, especially when the criterion of Murray and Washington was used. However, the differences were not significant ($P > 0.05$, chi-square test). When slides were prepared by selecting purulent flecks, 83% of the specimens showed complete agreement of results among triplicate slides, whereas 93% of the mixed samples were reproducible. The results obtained by the two interpretive criteria were 86 and 91% reproducible, respectively.

Cultures of all six samples from each of 20 different specimens revealed essentially identical results. The predominant microorganisms in 18 of 20 specimens were identical in all six samples. *Streptococcus pneumoniae* was found in two of three mixed samples and in one of three selected samples from one specimen. In the other specimen, five of the six samples yielded predominant growth of *Streptococcus pneumoniae*, and one selected fleck was negative for *Streptococcus pneumoniae*. Although the predominant microbial flora was generally recovered consistently, some variability was observed in the nonpredominant colonizing microorganisms that were recovered from different samples. Two specimens contained potentially pathogenic species as part of the nonpredominant microflora of some but not all of the cultures. One contained an unidentified lactose-nonfermenting, gram-negative bacillus (1+ to 2+ amounts) in all of the cultures except for one mixed sample. Another specimen contained 1+ to 2+ growth of *Haemophilus* sp. in two of three selected samples and two of three mixed samples.

DISCUSSION

An expectorated sputum specimen which contains a potential pathogen presents a diagnostic dilemma to the physician. It is often impossible to determine whether the potential pathogen is an etiological agent or represents oropharyngeal contamination. The amount of oropharyngeal contamination can be judged by evaluating the relative number of squamous epithelial cells in the specimens. Those specimens that are obviously contaminated are less likely to yield interpretable results, although potential pathogens may be present. Correlation of data obtained from expectorated sputa and transtracheal aspirates should provide meaningful insight into the interpretive problem (4, 7, 8). However, because of the risk inherent in collecting transtracheal aspirates, the patients who are examined in this way bias the data.

By the six different published criteria that were compared in this study, specimens with two to three predominant species were generally acceptable, and those with three to four predominant species were unacceptable. About 40% of the acceptable specimens contained potential pathogens as predominant or copredominant members of the microbiota, whereas 19 to 30% of the unacceptable specimens contained potential pathogens. The significance of potential pathogens in specimens which show excessive oropharyngeal contamination cannot be determined from our data. We only wished to compare the relative efficiency of the six different screening criteria.

The criteria described by Geckler et al. (4), Murray and Washington (7), and Van Scoy (8) are the simplest to describe—reject specimens with more than a defined number of cells per LPF. However, variations in the thickness of material in different areas of the slide may require extensive examination to obtain an overall average for each slide. For minimizing this variable, the other criteria involve assessment of the ratio of inflammatory cells to squamous epithelial cells in several areas of the slide. The authors of these criteria assume that both cell types will be distributed over the slide in similar proportions and thus that a ratio is likely to be more reproducible. We feel that such an assumption is valid, but we did encounter some difficulties in assessing ratios when there were low numbers of both types of cells on a slide. Bartlett's criterion also takes the presence of mucus into account. Murray and Washington (7) reported that 90% of their specimens contained

mucus; we observed mucus in only 68% of our specimens. Failure to observe mucus on a gram-stained smear would increase the rejection rate, since mucus adds a positive value to the Q score.

Murray and Washington (7) initially reported that 45% of their specimens were rejected. Barry and Rowley (unpublished data) initially observed a 35% rejection rate. In the present report, those two methods found a lower proportion of specimens unacceptable (25 and 23%, respectively). The current rejection rate might be explained by the improved quality of specimens being submitted to our laboratory during the study period. Heineman and Radano (6) reported a rejection rate of 32%, comparable to our 28% rate.

Ideally, a screening method should disqualify a high proportion of specimens without clinically relevant microorganisms yet not reject many specimens which contain potential pathogens. Heineman and Radano (6) suggested that a cost-effective system should reject at least 22% of the specimens that are received: if more than 78% of the specimens were accepted, it would be more efficient to culture all specimens without screening. That concept does not consider the potentially misleading information that can be derived from cultures of obviously unsatisfactory specimens. In any case, a Gram stain should be performed and the cellular composition of the specimen described so that the physician can interpret the results. The cost of examining the Gram stain is the same; screening only involves prompt examination of the slides and rejection of obviously unsatisfactory specimens.

Of the six criteria compared in this report, the methods of Murray and Washington (7) and of Barry (1) reached a reasonable compromise, finding 25 and 23%, respectively, of the samples unacceptable without rejecting many potentially significant cultures, i.e., 13% of the specimens with potential pathogens were rejected under both methods. In our hands, the fewest specimens (9%) were rejected by the criterion of Geckler et al. (4), and the most specimens (28 and 29%, respectively) were rejected by the criteria of Heineman and Radano (6) and Van Scoy (8).

The method of Murray and Washington depends only on the estimation of squamous epithelial cells, whereas that of Barry estimates the ratio of white blood cells to squamous epithelial cells. The question of which method is most reliable remains to be determined. Our studies failed to demonstrate any statistically significant differences in the reproducibility of the two methods. Although the method of Murray and Washington tended to give somewhat more reproducible results than did that of Barry, a more extensive study would be required to determine whether the minor differences that we noted are truly significant. In the interim, we concluded that either method can be used for screening expectorated sputum specimens. The four other published criteria might be preferred for reasons not considered in this review. The current study clearly documents the fact that the six different screening criteria yield similar but not identical results.

LITERATURE CITED

1. **Barry, A. L.** 1978. Clinical specimens for microbiologic examination, p. 92–96. *In* P. D. Hoeprich (ed.), Infectious diseases, 2nd ed. Harper & Rowe, Publishers, New York.
2. **Bartlett, J. G., N. S. Brewer, and K. J. Ryan.** 1978. Cumitech 7, Laboratory diagnosis of lower respiratory tract infections. Coordinating ed., J. A. Washington II. American Society for Microbiology, Washington, D.C.
3. **Bartlett, R. C.** 1974. Medical microbiology: quality, cost and clinical relevance, p. 24–31. John Wiley & Sons, New York.
4. **Geckler, R. W., D. H. Gremillion, C. K. McAllister, and E. Ellenbogen.** 1977. Microscopic and bacteriological comparison of paired sputa and transtracheal aspirates. J. Clin. Microbiol. **6**:396–399.
5. **Heineman, H. S., J. K. Chawla, and W. M. Lofton.** 1977. Misinformation from sputum cultures without microscopic examination. J. Clin. Microbiol. **6**:518–527.
6. **Heineman, H. S., and R. R. Radano.** 1979. Acceptability and cost savings of selective sputum microbiology in a community teaching hospital. J. Clin. Microbiol. **10**:567–573.
7. **Murray, P. R., and J. A. Washington II.** 1975. Microscopic and bacteriologic analysis of expectorated sputum. Mayo Clin. Proc. **50**:339–344.
8. **Van Scoy, R. E.** 1977. Bacterial sputum cultures, a clinician's viewpoint. Mayo Clin. Proc. **52**:39–41.