Utility of Gram Staining for Evaluation of the Quality of Cystic Fibrosis Sputum Samples

Bindu Nair, Jenny Stapp, Lynn Stapp, Linda Bugni, Jill Van Dalfsen, and Jane L. Burns

Division of Infectious Disease, Department of Pediatrics, University of Washington, Cystic Fibrosis Foundation Therapeutic Development Network Core Microbiology Laboratory, and Chiron Corporation, Seattle, Washington

Received 28 December 2001/Returned for modification 18 March 2002/Accepted 12 May 2002

The microscopic examination of Gram-stained sputum specimens is very helpful in the evaluation of patients with community-acquired pneumonia and has also been recommended for use in cystic fibrosis (CF) patients. This study was undertaken to evaluate that recommendation. One hundred one sputum samples from CF patients were cultured for gram-negative bacilli and examined by Gram staining for both sputum adequacy (using the quality [Q] score) and bacterial morphology. Subjective evaluation of adequacy was also performed and categorized. Based on Q score evaluation, 41% of the samples would have been rejected despite a subjective appearance of purulence. Only three of these rejected samples were culture negative for gram-negative CF pathogens. Correlation between culture results and quantitative Gram stain examination was also poor. These data suggest that subjective evaluation combined with comprehensive bacteriology is superior to Gram staining in identifying pathogens in CF sputum.

Gram staining is considered key in the workup of sputum from patients with pneumonia and other lower respiratory tract infections (10). Unfortunately, sputum samples are often contaminated with saliva and resident oropharyngeal microbial flora. This contamination can be quantified by evaluating Gram-stained sputum samples for polymorphonuclear leukocytes and squamous epithelial cells. The resulting objective measure, called the quality (Q) score, has been validated and recommended for use in evaluating the quality of the specimen (1).

Cystic fibrosis (CF) lung infections are chronic and polymicrobial, occasionally marked by exacerbations accompanied by increased, purulent sputum production and systemic symptoms. Antibiotic treatment based on sputum culture results is a mainstay of CF management. The specific use of the Gram stain for evaluating the adequacy of sputum samples from patients with CF has also been recommended (4). In addition, past literature reports excellent sensitivity, specificity, and predictive values of examination of the morphology of bacteria seen in Gram-stained specimens compared with microbiological culture techniques for the prediction of specific organisms in CF sputum samples (11).

At the Therapeutic Development Network (TDN) Core Microbiology Laboratory, we have had experience processing thousands of CF respiratory samples. It is the impression of the technologists at the bench that the Q score of the Gram-stained specimen is less useful than subjective evaluation of the specimens for adequacy. In addition, because specimens are solubilized and diluted prior to plating on selective media (2), it is rare that organisms are detected by Gram staining and not cultured from the sample. To validate these impressions, we performed a prospective study to compare the objective scoring (Q score) and Gram stain evaluation with subjective characterization of purulence and quantitative culture results.

MATERIALS AND METHODS

Study design. The present study was a substudy in the investigation of tobramycin-containing media for the isolation of gram-negative CF pathogens (13). Patients with a diagnosis of CF and a history of being culture positive for Pseudomonas aeruginosa in the lower respiratory tract were enrolled from seven CF centers across the United States. To be eligible, all patients were required to provide informed consent, to be able to produce sputum, and to have a culture history of P. aeruginosa in the lower airway and, if currently receiving aerosolized antibiotics, the specimen was to have been collected at least 8 h after the last dose. Each patient provided a single expectorated sputum sample of at least 0.5 ml that was immediately placed on ice or at refrigeration temperature and shipped on ice to the TDN Core Microbiology Laboratory at Children's Hospital and Regional Medical Center in Seattle, Wash. All samples were received by the laboratory within 48 h of expectoration.

Microbiology methods. Samples were evaluated by gross appearance and subjectively categorized as purulent (containing pus with or without mucus, saliva, or blood) or nonpurulent (no pus present), with subcategories of mucus (mucus strands present), watery (saliva present), or bloody (blood present). These descriptions are in daily use in both the TDN Core Microbiology Laboratory and the CF clinical microbiology laboratory and are used to characterize all CF sputum samples received. The individuals performing the subjective categorization were specifically trained by the authors to evaluate samples according to the above-mentioned criteria, and their ability to do so accurately was verified.

Gram stains were prepared according to standard procedure (3), with the caveat that only the purulent-appearing portions of the sputa were sampled. Each Gram-stained specimen was read twice in a blinded fashion by two medical microbiology technologists (L. S. and L. B.), each with over 20 years of clinical laboratory experience, and scored for polymorphonuclear leukocytes and squamous epithelial cells. The quality of the sample was rated by the Q score, assigning a value to the number of neutrophils and squamous epithelial cells per low-powered field (10× lens objective) and summing the results. The potential range was 0 to +3 for neutrophils and −3 to 0 for squamous epithelial cells, resulting in a potential range of values for the Q score from −3 to +3 (see reference 8 for details of scoring). The results from all four Q-score readings were averaged. In addition, each Gram stain was evaluated for the presence of gram-negative bacilli and the results were reconciled between the two evaluators.
TABLE 1. Association between subjective descriptions and Q score

<table>
<thead>
<tr>
<th>Descriptive term</th>
<th>No. of samples</th>
<th>Average Q score</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purulent</td>
<td>38</td>
<td>+1.1</td>
<td>−3−+3</td>
</tr>
<tr>
<td>Purulent-bloody</td>
<td>4</td>
<td>+1.4</td>
<td>0−+3</td>
</tr>
<tr>
<td>Purulent-mucus</td>
<td>47</td>
<td>+1.1</td>
<td>−1−+3</td>
</tr>
<tr>
<td>Purulent-watery</td>
<td>11</td>
<td>+0.6</td>
<td>−1−+2</td>
</tr>
<tr>
<td>Nonpurulent-mucus</td>
<td>1</td>
<td>−1</td>
<td>NA</td>
</tr>
</tbody>
</table>

* a n = 101.  
* b NA, not applicable.

For the purpose of this study, identification of gram-positive bacteria as well as *Haemophilus influenzae* was excluded.

Sputum samples were solubilized in dithiothreitol (Sputolytin; Calbiochem-Behring, La Jolla, Calif.), diluted, and set up for quantitative culture on selective media, as previously described (2), including MacConkey agar and oxidase fermentative basal media containing polymyxin, bacitracin, and lactose. In brief, bacteria quantitated as 4+, 3+, 2+, and 1+ had at least 10^6, 10^5, 10^4, and 10 CFU per ml, respectively. Exact quantitation of *P. aeruginosa* was performed and recorded. Semi-quantification of other potential gram-negative CF pathogens, specifically, *Burkholderia cepacia* complex, *Stenotrophomonas maltophilia*, and *Achromobacter xylosoxidans*, was also recorded.

**Statistical analysis.** Data analysis included using descriptive measures (frequencies and percentages) to summarize study variables. Comparisons of Q score and culture quantitation were performed by categorizing Q scores as above or below the recommended cutoff value of 1 (8) and culture quantitation as above or below a clinically defined value of 1 (10 CFU/ml) to calculate the positive and negative predictive values with binomial exact 95% confidence intervals (CIs). Linear regression analysis was performed to determine whether Gram stain quantitation was predictive of culture quantitation.

**RESULTS**

**Demographics of study subjects.** Sputum samples were recovered from 101 CF patients. The ages of these patients ranged from 6 to 57 years (mean age, 20.9 years), with approximately equal numbers of males and females included in the study. Only gram-negative organisms that were potentially pathogenic in CF were evaluated in this study. Eighty-six patients yielded *P. aeruginosa* cultures, 4 yielded *B. cepacia* complex cultures, 9 yielded *S. maltophilia* cultures, and 10 yielded *A. xylosoxidans* cultures.

**Association between subjective evaluation and Q score.** Of the 101 samples examined, all but one was subjectively determined to be purulent (Table 1). The Q score of the nonpurulent sample was determined upon all four evaluations to be −1. Average Q scores were computed for each of the purulent categories. For these four categories, there was a wide range of Q scores and a similar average Q score. For the purulent, purulent-bloody, and purulent-mucus categories, the average Q scores were determined to be greater than +1, while the average Q score was +0.6 for the purulent-watery category.

**Association between bacterial density, Q scores, and subjective measures of purulence.** There was one culture that was inadvertently quantitatively cultured for only *P. aeruginosa* and *B. cepacia* and not other gram-negative pathogens, thereby excluding that sample from the analyses of association between bacterial density and measures of purulence or Gram stain phenotype (*n* = 100 for these analyses). There were 11 specimens from which no gram-negative bacteria were recovered by microbiologic culture techniques. Of these, nine were described as purulent, one was described as purulent-mucus, and one was described as purulent-watery. The Q scores for these samples ranged from −2.25 to +2, with a mean of 0.95. Of the remaining 89 specimens, there were 70 with 4+ bacteria, 13 with 3+ bacteria, 5 with 2+ bacteria, and 1 with 1+ bacteria. A Q score of ≥1 is the standard cutoff for acceptance of a sample for culture (8). The positive predictive value of this cutoff value with culture quantitation (≤1 versus ≥2+) as a gold standard was 84.7% (95% CIs, 73.0 to 92.8%). The negative predictive value was 98.8% (95% CIs, 2.7 to 23.1%). By using the Q score cutoff of ≥1 for acceptance, only 59 of the 100 samples would have been accepted for microbiologic culture. Of the 41 rejected samples, all but four (20%) were culture positive for gram-negative CF pathogens at a density of ≥2+. All 100 samples were ranked by average Q score (mean of all 4 determinations) and examined for semiquantitative culture results (Table 2); no correlation could be identified. Of the full range of possible average Q scores, the percentage of samples with culture quantitations of 4+ was similar among all categories with adequate numbers for analysis (range, 68 to 72.4%).

**Association of culture results with Gram stain phenotype.** The sputum specimens were Gram stained and microscopically examined for the presence and numbers of gram-negative bacilli (Table 3). Linear regression analysis identified a poor correlation between Gram stain quantitation and culture quantitation (r = 0.21). Microscopic identification of abundant or moderate numbers of gram-negative bacilli correlated with the culture density of >1+ organisms 94% of the time. However, 71 and 48% of the specimens microscopically identified as having few and no gram-negative bacilli, respectively, quantitatively cultured 4+ bacteria. There were 15 Gram stains with bipolar-stained gram-negative bacilli that exhibited what is described as the classic morphology for *B. cepacia* (11), only two of which were culture positive for *B. cepacia* complex organisms. Two additional specimens that cultured positive for *B.

**TABLE 2. Association between Q score and culture density**

<table>
<thead>
<tr>
<th>Average Q score</th>
<th>No. of samples</th>
<th>Culture density (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4+</td>
<td>3+</td>
</tr>
<tr>
<td>≥2</td>
<td>25</td>
<td>17 (68)</td>
</tr>
<tr>
<td>1–1.75</td>
<td>34</td>
<td>24 (70.6)</td>
</tr>
<tr>
<td>0–0.75</td>
<td>29</td>
<td>21 (72.4)</td>
</tr>
<tr>
<td>−1–−0.25</td>
<td>10</td>
<td>7 (70)</td>
</tr>
<tr>
<td>−0.25–0</td>
<td>22</td>
<td>1 (50)</td>
</tr>
<tr>
<td>≤−0.25</td>
<td>2</td>
<td>1 (50)</td>
</tr>
</tbody>
</table>

* Only gram-negative rods were quantified (n = 100).

**TABLE 3. Association between quantitation by Gram stain and culture density**

<table>
<thead>
<tr>
<th>Gram stain quantitation (organisms/field)</th>
<th>No. of samples</th>
<th>Culture density (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4+</td>
<td>3+</td>
</tr>
<tr>
<td>Abundant (≥30)</td>
<td>21</td>
<td>17 (80.9)</td>
</tr>
<tr>
<td>Moderate (6–30)</td>
<td>14</td>
<td>13 (92.9)</td>
</tr>
<tr>
<td>Few (0–5)</td>
<td>38</td>
<td>27 (71.1)</td>
</tr>
<tr>
<td>None</td>
<td>27</td>
<td>13 (48.1)</td>
</tr>
</tbody>
</table>

* Only gram-negative rods were quantified (n = 100).

* Field = oil immersion field.

* Two organisms were quantified as 3+.
**DISCUSSION**

Sputum cultures are an inherently difficult sample from which to determine the microbiological etiology of pneumonia because of the high potential for contamination from oral flora. Gram staining to determine the adequacy of a sample has proven quite useful in making a microbiological diagnosis of acute bacterial pneumonia (5, 10, 15). Rejection of samples because of what appears to be upper airway contamination (the presence of squamous epithelial cells and a lack of polymorphonuclear leukocytes) is appropriate for expectorated sputum samples in patients with pneumonia. In one study, 60% of submitted samples were contaminated with saliva and misinformation about culture results would have been reported without microscopic examination (6).

The use of sputum culture is critical in determining the etiology of pulmonary exacerbation in CF, and Gram staining has been recommended to ensure the adequacy of CF samples (4). However, in CF it is rare to see a nonpurulent sample from patients capable of expectorating. CF patients often have neutrophilic infiltrates and large numbers of bacteria per gram of sputum (14). In addition, the organisms that are usually reported in CF sputa are distinct from normal oropharyngeal flora (2). The presence of *P. aeruginosa* in oropharyngeal culture from a CF patient is predictive of a positive culture from the lower airway (9). Thus, it is unlikely that a culture result for a CF sputum specimen that appears grossly purulent is not representative of lower airway infection. This hypothesis has been confirmed in the current study. All but 1 of the 101 sputum samples submitted for culture were subjectively described as purulent and would not have been rejected based on subjective criteria. However, by using the Q score from an objective measurement of the quality by Gram stain, 41% of the samples would have been rejected. Of the samples that would have been rejected based on Q scores alone, the vast majority grew large numbers of gram-negative bacilli, known to be pathogenic in CF. Thus, this objective measure does not appear to be useful for sputum samples from patients with CF.

There are a number of potential explanations for this lack of association between the Gram stain Q score, subjective categorization, and culture results. The most likely of these is the rheological character of CF sputum. Because of the high concentration of DNA and mucin, CF sputum can be viscous and cohesive, making it difficult to obtain a uniform specimen layer. This results in an uneven distribution of polymophonuclear leukocytes, squamous epithelial cells, and bacteria. Even with directed sampling of the most purulent portion of each specimen, the stickiness of the sputum may result in nonuniformity on a smear. However, cultures in our laboratory are performed using a mucolytic agent (dithiothreitol) with subsequent vigorous mixing and dilution. This method is likely to release organisms trapped within the mucus matrix, improving the yield on culture. This has been demonstrated previously in our laboratory (2). Another potential reason for the lack of association of Q scores with the results of subjective evaluation may lie in the method by which the Q score is calculated. Both neutrophils and squamous epithelial cells are quantified, and the relative numbers of each are used in the calculation. In CF, the inflammation present in the airways results in large numbers of neutrophils in the sputum. In addition, there is some evidence of squamous metaplasia in CF (12). Thus, there are likely to be relatively more squamous epithelial cells as well, affecting the ratio used for calculation of the Q score.

The lack of inter- and intrareader reliability of the Q score, meant to be an objective measurement, is also a point of concern. Only 21 of the 101 samples were scored identically by two very experienced readers when performed in duplicate. This, too, may be a result of the lack of homogeneity in sampling. Since the same fields are examined each time a slide is read, variability in the uniformity of the smear results in variability in the reading.

The use of the Gram stain morphology to predict likely pathogens in CF sputum was also examined. There was good correlation between the presence of large numbers of gram-negative bacilli on Gram stains and positive cultures for *P. aeruginosa*. On the other hand, the absence of gram-negative bacilli or the presence of only a few of these bacilli on Gram-stained specimens correlated poorly with culture results. As noted above, this likely results from the lack of homogeneity of the smear caused by an inability to adequately sample the entire specimen. Similar results were reported for *B. cepacia*, but the numbers were very small. Of potential concern are the 15 samples that had what is reported in the literature as the classic Gram stain appearance for *B. cepacia* (11) but did not grow the organism on culture. If culturing (on selective media with solubilization and dilution) were considered the gold standard for a microbiological diagnosis, these samples would represent false positives. In fact, other organisms might be observed as gram-negative bipolar-stained bacilli, including *Bacteroides* species, which are part of the normal mouth flora (3). Additionally, it is difficult to use Gram staining to distinguish *Burkholderia* from *Stenotrophomonas, Ralstonia*, and other miscellaneous species of gram-negative bacteria. While it is possible that Gram staining may be more sensitive than microbiologic culturing in identifying *B. cepacia*, examination of additional samples by molecular techniques such as PCR would be necessary to evaluate this possibility. It is more likely that, with the large number of genomovars now included in the *B. cepacia* complex, the bipolar staining described as classic is not present in all genomovars. The design of the present study, culturing samples from patients at seven different CF centers throughout the United States, is more likely to result in *B. cepacia* complex organisms representing multiple genomovars.

The culture findings in CF are distinct from those in patients with pneumonia and chronic bronchitis. In patients with pneumonia, the most common bacterial pathogens are organisms...
such as *Streptococcus pneumoniae* and *H. influenzae*, which may be normal oropharyngeal flora. Thus, salivary contamination of a specimen might yield those organisms on culture. For this reason, sputum cultures directed by Gram stain results are recommended by some investigators (6). In patients with chronic bronchitis, microscopic examination of the Gram stain was found to be superior to culture of sputum samples (7), again because of the possibility of encountering normal oropharyngeal flora in inadequately obtained samples. Lung infections in patients with CF are caused by a much more distinctive flora for which a range of selective media are recommended for routine use (4).

The cost of performing a Gram stain on sputum is relatively low compared with the cost of medical care for CF patients. However, one must also account for the time, expense, and patient discomfort involved in obtaining a new sample if the first is inappropriately rejected based on the Q score. This would significantly increase the cost of performance of the Gram stain.

The potential limitations of this study include the fact that samples were shipped on ice from distant sites to arrive within 48 h, which may have resulted in changes in bacterial morphology affecting phenotypic identification or disintegration of polymorphonuclear leukocytes impacting the Q score. In addition, all patients were required to have a culture history of *P. aeruginosa* in the lower airway, thus biasing the results to those individuals with mostly purulent sputum samples. Also, because the culture conditions of the parent study were designed to retrieve only gram-negative bacteria, the presence of gram-positive bacteria that may have been the cause of sample purulence could very well have been overlooked. However, the former two factors may be seen clinically because of delays in processing and would thus support the abandonment of the Gram stain for the processing of CF sputum samples. And the latter two factors which might be expected to increase the purulence of specimens would likely make the Q score more, rather than less, predictive.

The results presented here suggest that subjective evaluation of sputum samples from patients with CF is superior to the Gram stain in identifying adequacy for culture. In addition, evaluation of bacterial morphology does not add significantly to the detection of organisms when compared with recommended culturing techniques. If semiquantitative culture is performed using recommended selective media and specimen solubilization (2, 4), the use of the Gram stain does not necessarily aid in the detection of gram-negative pathogens in CF sputum.

**ACKNOWLEDGMENTS**

This study was sponsored by Chiron Corporation, Seattle, Wash. Statistical consultation was provided by Julia Emerson. We thank the following principal investigators (PI) and study coordinators for their support of this study: H. Eigen (PI), M. Blagburn, and D. Terrill, Indiana University and Medical Center; R. Anbar (PI), D. Lindner, and L. Grabowski, SUNY Health Science Center; S. Nasr (PI) and E. Sakmar, University of Michigan; C. Ren (PI) and L. Maftia, University Hospital, Children’s Medical Center at Stonybrook; R. Gibson (PI), S. McNamara, and P. Joy, Children’s Hospital and Regional Medical Center, Seattle, Wash.; P. Hiatt (PI), J. Morris, and D. Trecece, Baylor College of Medicine/Texas Children’s Hospital; M. McCarty (PI) and D. Harrington, Deaconess Medical Center, CF Clinic, Spokane, Wash. Most importantly, we thank the patients who volunteered to participate in the study.

**REFERENCES**