Interpreting the Sputum Gram Stain Report

In 1884 Hans Christian Joachim Gram, a Danish physician, developed the Gram stain.1 This differential stain facilitates observation and subsequent characterization of bacteria according to their shape, size, group, and staining reaction. The Gram stain is the most important and widely used stain in the fields of clinical microbiology, infectious diseases, and infection control.

Gram stains and cultures of sputum specimens are performed to detect potential respiratory pathogens. Gram stain results are available much sooner than culture results, and help physicians choose empiric therapy for patients with clinical signs of acute bacterial pneumonia.2 However, the usefulness of the sputum Gram stain is controversial because the results are not always reliable, culture results can be ambiguous, and pathogens do not always grow as expected.3-8

Determining potential respiratory pathogens from a Gram stain report prior to availability of culture results can be difficult. Problems can arise if Gram stains are misread or misinterpreted, specimens do not accurately reflect materials from the lower respiratory tract, the Gram stain report does not reflect subsequent culture results, or the Gram stain or culture do not reflect the actual cause of the illness.

This study focuses only on the relationship of Gram stain to culture. Investigators have noted relationships between Gram stains and cultures of Streptococcus pneumoniae,6-10 Haemophilus influenzae,6,11,12 and Moraxella catarrhalis.13,14 However, few, if any, studies address how physicians should interpret sputum Gram stain results in general (ie, for all isolates cultured routinely). In this study, statistical relationships (sensitivity, specificity, positive and negative predictive values, and positive likelihood ratio) between some common Gram stain morphotypes and culture results were used to devise guidelines for interpreting the Gram stain report.

ABSTRACT A retrospective review of patient records was performed to determine the statistical relationship between sputum Gram stain and culture results in adult men at a Veterans Affairs Medical Center in the Midwest. Gram stain results from 1996 were statistically compared with culture results for 2,105 sputum specimens. Positive predictive values for common Gram stain morphotypes and their corresponding organisms varied from 7.1% to 90.6%. In some instances, the absence of a particular Gram stain morphotype (eg, gram-negative diplococci) was more predictive of the absence of an organism in culture (eg, Moraxella catarrhalis) than its presence. Also, Gram stain reports noting the presence of gram-negative bacilli were not predictive of cultures with gram-negative bacilli potential respiratory pathogens (positive predictive values 32.4%-54.9%). In conclusion, Gram stain results often were not accurate predictors of sputum culture results. One way to improve agreement between Gram stain and culture results is to develop Gram stain interpretation guidelines based on statistical relationships between stain and culture.

Materials and Methods

The study was performed in a 389-bed, tertiary-care, general medical and surgical teaching institution primarily serving male veterans in Oklahoma and north central Texas. The study period was 1996, during which 7,405 inpatients and 263,489 outpatients were seen at the hospital and 2,105 sputum cultures and Gram stains were performed.

Sputum Culture and Gram Stain

A routine bacterial sputum culture included a Gram stain.15 The Gram stain report noted sputum quality, and presence of neutrophils and bacterial morphotypes (eg, gram-positive cocci in pairs and clusters) semiquantitated. A positive Gram stain result indicated specific bacterial

From the Veterans Affairs Medical Center, and Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City.

Reprint requests to Dr Flournoy, VAMC (113), 921 NE 13th St, Oklahoma City, OK 73104-5007.
Fig 1. Positive sputum specimen showing moderate number of gram-positive cocci in pairs (Gram stain, original magnification ×1,000).

Table 1. Criteria for Grading the Quality of Sputum Specimens

<table>
<thead>
<tr>
<th>Specimen Grade</th>
<th>Squamous Epithelial Cells/lpf</th>
<th>Neutrophils/lpf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>0–10</td>
<td>NA (usually &gt;25)</td>
</tr>
<tr>
<td>Fair</td>
<td>11–19</td>
<td>NA (usually &gt;25)</td>
</tr>
<tr>
<td>Poor</td>
<td>&gt;19</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Inadequate</td>
<td>&gt;19</td>
<td>&lt;10</td>
</tr>
</tbody>
</table>

Lpf indicates low power field (10x ocular with 10x objective); NA, not applicable.

Fig 2. “Fair” sputum specimen—11 to 19 squamous epithelial cells per low power field (Gram stain, original magnification ×100).

Fig 3. “Good” sputum specimen—0 to 10 squamous epithelial cells per low power field (Gram stain, original magnification ×100).

morphotypes or mixed oral flora (Fig 1); a negative result indicated the absence of microorganisms (ie, no organisms seen).

The Gram stain was usually read within several hours after the specimen was received in the laboratory, and the culture was initially read the following morning. Specimens were cultured in 100-mm-diameter plates containing 5% sheep blood, chocolate, and MacConkey agars, and incubated in 5% carbon dioxide at 35°C for 24 hours after inoculation. Some plates were then held longer, depending on microbial flora present and the need to identify and semiquantitate isolates and perform antimicrobial susceptibility tests.

As shown in Table 1, sputum specimens were graded “good” if they had 10 or fewer squamous epithelial cells per low-power field (100 × total magnification), “fair” with 11 to 19 squamous epithelial cells, and “poor” with more than 19 squamous epithelial cells and more than 10 WBC per low-power field, based on previously described criteria.16,17 Low-power magnification was used to detect and quantitate squamous epithelial cells and WBC (Figs 2 & 3); however, microorganisms were observed under oil immersion (1,000 × total magnification). Approximately 10 to 20 low-power fields were scanned
before determining the specimen’s grade. Because the established grading system potentially underestimates contamination from common oral flora, the poorer grade was assigned if a borderline result was obtained.

“Inadequate” sputum specimens (>19 squamous epithelial cells and <10 WBC per low-power field) are not routinely cultured and were not included in this study (Fig 4). Poor specimens often have mixed oral flora that were counted as positive. The value of Gram stain data from poor specimens was reduced because there is a greater likelihood of contamination by oral flora.

Positive cultures grew potential respiratory pathogens (gram-negative bacilli, S pneumoniae, Streptococcus pyogenes, other β-streptococci, M catarrhalis, and Staphylococcus aureus); negative cultures did not grow potential respiratory pathogens. Sputum cultures with fewer than 5 colonies per plate of potential respiratory pathogens were counted as having only common oral flora.

**Statistical Analyses**

Statistical analyses (sensitivity, specificity, positive, and negative predictive values, and positive likelihood ratio) were performed with published formulas. Likelihood ratios greater than 10 or less than 0.1 alter probability greatly, ratios from 5 to 10 and 0.1 to 0.2 alter probability moderately, 2 to 5 and 0.5 to 0.2 minimally, and 1 to 2 and 0.5 to 1 rarely. Differences in proportions were determined with the χ² test, and P < .05 was considered statistically significant (1 df).

**Quality Control**

During the study our laboratory was enrolled in the College of American Pathologists Bacteriology Survey, which tested the ability of personnel to accurately identify bacteria and perform Gram stains. Our personnel successfully identified microorganisms and interpreted Gram stains in a blind fashion during the study period.

**Results**

Data on Gram stains, culture results, and sputum quality are shown in Table 2. Poor sputum samples were more likely to yield positive Gram stain results and negative cultures than good to fair samples (P < .05). Sputum samples graded good were more likely to yield Gram stain and culture results that agreed (ie, both culture and Gram stain were positive or both were negative), and were more likely to yield positive culture results than were fair or poor specimens (P < .05). Also, good sputum specimens had more potential respiratory pathogens per culture than did specimens of lesser quality.
Table 3. Statistical Comparison of Gram Stain and Culture Results

<table>
<thead>
<tr>
<th>Gram Stain</th>
<th>Culture</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Positive Likelihood Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gpc in clusters</td>
<td>Staphylococci</td>
<td>37.1</td>
<td>99.5</td>
<td>90.6</td>
<td>91.9</td>
<td>74.2</td>
</tr>
<tr>
<td>Gpc in clusters†</td>
<td>Staphylococcus aureus</td>
<td>36.1</td>
<td>91.2</td>
<td>30.3</td>
<td>93.1</td>
<td>4.1</td>
</tr>
<tr>
<td>Gpc in chains</td>
<td>Streptococci</td>
<td>48.3</td>
<td>97.4</td>
<td>58.4</td>
<td>96.1</td>
<td>18.6</td>
</tr>
<tr>
<td>Gpc in pairs‡</td>
<td>Streptococcus pneumonia</td>
<td>92.2</td>
<td>54.7</td>
<td>7.1</td>
<td>99.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Gndc</td>
<td>Moraxella catarrhalis</td>
<td>93.1</td>
<td>83.8</td>
<td>7.4</td>
<td>99.9</td>
<td>5.7</td>
</tr>
<tr>
<td>Small or tiny Gnb</td>
<td>Haemophili</td>
<td>26.0</td>
<td>99.3</td>
<td>81.4</td>
<td>92.1</td>
<td>37.1</td>
</tr>
<tr>
<td>Gnb—all sputa</td>
<td>Gnb</td>
<td>74.9</td>
<td>60.2</td>
<td>43.9</td>
<td>85.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Gnb—good sputa</td>
<td>Gnb</td>
<td>75.6</td>
<td>59.6</td>
<td>54.9</td>
<td>78.9</td>
<td>1.8</td>
</tr>
<tr>
<td>Gnb—fair sputa</td>
<td>Gnb</td>
<td>79.7</td>
<td>52.7</td>
<td>38.4</td>
<td>87.6</td>
<td>1.7</td>
</tr>
<tr>
<td>Gnb—poor sputa</td>
<td>Gnb</td>
<td>43.8</td>
<td>71.7</td>
<td>32.4</td>
<td>80.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

PPV indicates positive predictive value; NPV, negative predictive value; Gpc, gram-positive cocci; Gndc, gram-negative diplococci; Gnb, gram-negative bacilli.

*The number of cultures analyzed was 2,105 for each category. Of the 2,105 cultures, 20% were from outpatients, 36% from general inpatients, and 44% from patients in intensive care.
†A comparison was made between a Gram stain result of gram-positive cocci in clusters or clusters and pairs (with no mention of chains) and a culture result yielding staphylococci.
‡A comparison was made between a Gram stain result of gram-positive cocci in pairs or pairs and chains (with no mention of clusters) and a culture result yielding Streptococcus pneumoniae.

Statistical comparison of Gram stain and culture results is presented in Table 3. As anticipated, several Gram stain report comments or morphotypes were not predictive of culture results. For example, gram-negative diplococci did not predict for culture of *M catarrhalis* (positive predictive value 7.4%). Also, the observation of gram-negative bacilli on Gram stain did not predict the presence of gram-negative bacilli potential respiratory pathogens in culture, regardless of sputum specimen quality (positive predictive value 32.4%-54.9%). Positive likelihood ratios greater than 10 were considered most significant, occurring with gram-positive cocci in clusters and staphylococci, and gram-positive cocci in chains and streptococci, and small or tiny gram-negative bacilli and haemophili.

**Discussion**

The first step in obtaining useful Gram stain and culture results is procurement of an acceptable specimen. Collecting an expectorated sputum specimen usually includes an explanation to the patient, brushing the patient's teeth and/or rinsing the mouth, followed by having the patient eject an early morning bolus (from a deep cough) into a sterile container. Obtaining a good specimen might not be as easy as the literature suggests, because some patients (eg, those with congestive heart failure) are more likely to yield poor specimens, rinsing the mouth before expectoration does not always yield an acceptable specimen, and it can be difficult to macroscopically differentiate acceptable from nonacceptable sputum specimens at the bedside. Nevertheless, after the specimen is procured, Gram stained, and cultured, the physician must decide how to use the results. Generally I recommend that information generated from poor specimens be considered of less value than information from good or fair specimens. Ideally, new specimens should be collected to replace poor specimens. However, sometimes a poor specimen is the only specimen that can be obtained, and the clinician must decide how to use results from this specimen. Poor specimens can yield useful information and should be evaluated carefully when there is no other choice.

A positive Gram stain followed by a negative culture can be misleading. In this situation organisms seen on Gram stain can prompt the initiation of inappropriate or toxic therapy and delay making an accurate etiologic diagnosis. Data from the study hospital show that poor sputum specimens are more likely to produce positive Gram stains and negative cultures, compared with good or fair specimens.

In this study, statistical analyses indicated that Gram stain results often were misleading or not...
predictive of culture results. Perhaps the most misleading finding was the presence of gram-negative bacilli on Gram stains that did not predict the presence of gram-negative bacilli potential respiratory pathogens on corresponding cultures (ie, positive predictive value 32.4%-54.9%; see Table 3). This discrepancy between Gram stain and culture results is important because clinicians often are interested in the presence of gram-negative bacilli potential respiratory pathogens in the lower respiratory tract. Gram-negative bacilli might be seen on Gram stain but not grown on culture when the organisms are anaerobes or fastidious. Also, the laboratorian reading the culture may not recognize or detect gram-negative bacilli potential respiratory pathogens, especially if they are overgrown by common oral flora. Although the presence of gram-negative bacilli on Gram stain poorly predicted the presence of gram-negative bacilli potential respiratory pathogens in culture, the absence of gram-negative bacilli on Gram stain was a fair predictor for the absence of gram-negative bacilli in culture.

With these problems in mind, I devised guidelines for interpreting sputum Gram stain reports (Table 4), based on data from Table 3. However, these suggestions may not apply to some hospitals owing to variables that could affect the report, such as patient population, unique institutional flora, microbiologic culture techniques, media and evaluation policies, and the way physicians and other health care providers evaluate Gram stain results. Other investigators should repeat this type of study in their own settings to determine applicability.

When Fine and colleagues27 compared Gram stain with culture and serology, their statistics for *S pneumoniae* and *H influenzae*, respectively, were sensitivity 86% and 80%, specificity 72% and 88%, positive predictive value 43% and 73%, and negative predictive value 95% and 92%. Their statistics were similar to mine, except a lower positive predictive value for *S pneumoniae* and a lower sensitivity for *H influenzae* were found. Although it is difficult to determine the exact nature of the discrepancies, variables between the two studies might include what constituted a positive Gram stain for *S pneumoniae* and *H influenzae*, and definition of positive culture. In the current study, blood culture and serologic tests were not included.
In 1996 Reed and colleagues performed a meta-analysis to evaluate the sensitivity and specificity of the sputum Gram stain in community-acquired pneumococcal pneumonia. Sensitivity ranged from 15% to 100%, and specificity from 11% to 100%. They suggested that practitioners using the Gram stain be taught that a positive test requires on average more than 10 organisms resembling pneumococcus per oil immersion field in purulent sputum. However, all S. pneumoniae organisms may not be lancet shaped; some might be observed as gram-positive cocci in pairs, chains, or even clusters. In 1975 Lorian and Atkinson noted that pneumococci can look like bacilli, and staphylococci like diplococci, when exposed to antimicrobial agents. Therefore morphologic features can vary with the local "environment." This morphologic variation is a good example of how difficult it can be to compare Gram stain studies, all with different definitions of what constitutes a positive Gram stain and culture.

Conclusion

The relationship of Gram stain to culture needs to be examined for each specimen type (eg, sputum, urine, wound, blood). Studies need to examine whether guidelines (see Table 4) improve the ability of physicians to make etiologic diagnoses. However, for now, if physicians can more accurately interpret sputum Gram stain results, they might be able to better choose empiric therapy and therefore reduce morbidity and mortality, especially in seriously ill patients.

References