Evaluating the Clinical Effectiveness of Amniotic Fluid Assays in Predicting Respiratory Distress Syndrome in the Neonate

Joseph S. Amenta,1 Steffanie C. Brocher,2 and Andrea L. Serenko-Aber2

Amniotic fluid phospholipids from 346 patients’ specimens were quantified and evaluated against the clinical outcome (i.e., respiratory distress syndrome or normal respiratory function). Concentrations of lecithin, sphingomyelin, phosphatidylglycerol, and the lecithin/sphingomyelin reflectance ratio were evaluated by ordered frequency distribution and stepwise discriminant function analysis. The lecithin/sphingomyelin ratio was the best single test for discriminating between respiratory distress syndrome and normal pulmonary function in the fetus, slightly superior to lecithin assay alone. A combination of lecithin/sphingomyelin ratio and lecithin concentration, however, appeared to optimize the discriminant function, although the clinical significance of this test combination remained marginal. High concentrations of phosphatidylglycerol were correlated with high concentrations of lecithin, and virtually ruled out respiratory distress syndrome. Absence of phosphatidylglycerol was not diagnostic. High concentrations of sphingomyelin increased the probability of respiratory distress syndrome. We suggest the following stepwise series of tests to optimize diagnosis: phosphatidylglycerol concentration, sphingomyelin concentration, and finally lecithin–sphingomyelin ratio.

Additional Keyphrases: stepwise discriminant function analysis · phosphatidylglycerol · lecithin/sphingomyelin ratio · fetal status · cutoff value

Various assays of amniotic fluid have been recommended for use in evaluating the pulmonary status of the fetus (1). Although L/S by thin-layer chromatography is the technique most commonly used (2, 3), measuring the actual concentrations of specific phospholipids in the amniotic fluid has also been proposed (4–6). The comparative effectiveness of each of these analytes in discriminating between RDS and normal pulmonary function in the neonate remains unknown. Perhaps a single one of them would provide the maximum discriminant information and the rest only redundant information. Alternatively, a combination of assays may be required to attain the maximum discriminant power inherent in these tests.

In our laboratory, we routinely quantify phosphatidylglycerol, lecithin, and sphingomyelin, as well as the L/S reflectance ratio in amniotic fluid (7). For the present study we examined the records of 346 patients and compared the values obtained for these tests with the clinical outcome, i.e., respiratory distress syndrome or normal respiratory function in the newborn. We evaluated the effectiveness of each individual test and series of test groups by discriminant function analysis. As a result of these evaluations, we propose a stepwise interpretative scheme to maximize the discriminant power of these assays.

Materials and Methods

Samples

We analyzed specimens of amniotic fluid from 346 patients within 4 h of receipt in the laboratory. The specimens were maintained at 4 °C until analysis. To evaluate the procedures as fully as possible under actual working conditions, we analyzed all specimens, including those contaminated with blood or meconium. About 10% of the specimens were obtained vaginally after spontaneous rupture of membranes.

Specimens were assayed for L/S by the one-dimensional thin-layer chromatographic procedure of Gluck et al. (2, 3, 8). On all plates we also included a series of standards, from which we obtained a quantitative estimate of L, S, and PG concentrations in each unknown (7). Each plate also contained at least one control specimen, to validate the quantitative analysis on each plate. To attain a consistent spot density for all phospholipids, we sprayed each plate with 9 mol/L sulfuric acid before charring (2, 7, 9).

Patient Population and Diagnoses

Data was collected prospectively. We included patients in the study only if they delivered within 72 h of our obtaining the amniotic fluid specimen. Approximately 55% of these patients underwent cesarean section. The diagnosis of RDS in the neonate was based on previously reported criteria (7, 10). However, we found significant differences among physicians in applying the diagnosis of transient tachypnea of the newborn to borderline cases with respiratory difficulty. For purposes of consistency in this study, we classified neonates with a discharge diagnosis of either TTN or RDS as having RDS if supplementary oxygen was required for >24 h after delivery and TTN if it was required for <24 h. We did not evaluate the severity of RDS.

Descriptive Statistics

To display the predictive power of each assay, we sorted the cases successively by each test result, in ascending
order. After each sort, we subdivided the data base into 10 successive groups, each containing approximately 10% of the population under study. The remaining cases were added to the last subgroup. For each subgroup we recorded the number of cases and calculated the percentage with RDS. We found 58 cases of RDS in the study group, so the decision level for classifying RDS was set between the 58th and 59th largest value for each analyte. Thus, effectiveness of each analyte or analyte group was measured against a common theoretical analyte that discriminated perfectly between RDS and normal pulmonary function.

Using this type of data presentation, we could make a first approximation of the discriminant power of each test: we first counted the total number of misclassified cases in each subgroup, summed the total misclassified cases for the entire sample, and finally calculated the percentage of misclassified cases for the particular test or test group. For purposes of comparison, an analyte that discriminates perfectly would show 0% misclassified cases by this method. Conversely, for a test that does not discriminate between disease and non-disease, the percentage of patients with disease in each subgroup would be close to the percentage of patients with disease for the entire sample. The expected percentage of misclassified cases for the non-effective test would be given by the expression:

\[ \%\text{MSCL} = \left( \frac{N \times D - D^2}{200N^2} \right) \]

where \%MSCL is the percentage of cases misclassified, D is the number of cases with disease, and N is the total number of cases in the sample.

Hypothesis Testing

We evaluated the discriminant power of each test and each group of tests by applying a stepwise discriminant analysis program (SPSS, Version H, Release 9.0) available at the University of Pittsburgh Computer Center. This procedure first selects the test providing the maximum discriminant power, then adds tests sequentially to the discriminant function until further test additions contribute no more to the discriminant power of the derived linear function. We used a minimum probability of 0.10 for entering a value to the discriminant function; i.e., a test was entered into the discriminant function only if the probability of contributing additional diagnostic information exceeded 90%.

By this method we could compare the discriminant power of any two or more tests, determine the significance of the observed differences in discriminant power, and, equally important, evaluate whether the individual tests, by contributing independent discriminant information, should be combined to achieve maximum discrimination between disease and non-disease.

Results

Evaluation of All Tests as Used on All Samples

A comparison of the L/S values and the lecithin concentrations is shown in Figure 1. Although both L/S and lecithin concentrations achieved effective discrimination between cases of RDS and pulmonary maturity, the decile distributions suggested that L/S was the more effective test (Table 1). The L/S ratio showed a misclassified rate of 14.4%, contrasting with 18.4% for lecithin concentration (Table 2). This result would be expected, in that we did not eliminate amniotic fluids contaminated with non-surfactant phospholipids from our study. Contaminating phospholipids would be expected to add significantly to the assayed lecithin concentration and to shift cases incorrectly to the mature range of lecithin concentrations. Because contaminating phospholipids, however, also contain substantial amounts of sphingomyelin, their effect on the L/S ratio would be expected to decrease the "mature" L/S ratio to <2.0 (11, 12). Of note, we observed a significant rate of RDS in the borderline L/S range of 2.0-2.8 and in the borderline lecithin range of 12-30 mg/L (Table 1).

PG alone, showing a misclassified rate of 26.6%, was clearly a poor discriminator for RDS (Tables 1, 2). For purposes of comparison, the expected misclassified rate for a test showing no discriminant power in this sample was 27.9%. We usually detected PG when lecithin values exceeded 30 mg/L, consistent with previous observations in our laboratory (7). These data, however, corroborated that, while low values for PG were meaningless, concentrations $\geq 0.5$ mg/L effectively ruled out RDS.

Sphingomyelin, as expected, showed no discriminant power, in effect acting as a negative control (Table 1). The 28.9% were misclassified, as expected for a test with no discriminant power (Table 2). Discriminant function analysis, however, indicated that there was a slightly higher probability of RDS in patients showing high concentrations of sphingomyelin.

Of the three tests—L/S, lecithin, and PG—the L/S showed the lowest misclassified rate and the highest F value, indicating that it discriminated best between RDS and mature pulmonary function in the fetus. We applied stepwise discriminant analysis to this group of tests, to test whether the other assays contributed discriminant information independent of L/S. We also tested a previously derived function (7) that combined both L/S and lecithin concentration in a defined maturity index, MX1:

\[ MX1 = \left( \frac{L/S}{2.2} \right) + \left( \frac{Lec/16}{0.5} \right) \]  

(1)

Although the misclassified rate did not appear to differ from the results we obtained with the L/S ratio alone,
The stepwise discriminant analysis did suggest that lecithin concentration made a small independent contribution to the discriminant (Tables 1, 2). PG did not add to the discriminant power of this combination. The derived linear function (MX2) determined by the stepwise analysis was:

$$MX2 = (L/S \cdot 0.389) + (Lec \cdot 0.0187) - 1.96$$  \hspace{1cm} (2)

Similarly, when we tested the combination of tests, MX1, L/S, and lecithin, only the derived function provided independent discriminant information. These observations indicated that information on a combination of L/S and lecithin concentration provided the optimum discrimination between RDS and mature pulmonary function, although the clinical value of this small statistical increment in discriminant power remained doubtful.

Evaluation of Tests as Done on a Subdivided Sample

These results suggested an alternative strategy: divide the sample into subgroups based on values for PG and S and then analyze the groups sequentially. Because PG values of 0.5 mg/L or greater were rarely associated with RDS, we could immediately and confidently predict pulmonary maturity in this group of cases (Table 3). Because it was evident that high concentrations of sphingomyelin were more frequently associated with RDS than would be expected on a chance basis alone (Tables 2, 3), we next analyzed separately all remaining cases with sphingomyelin concentrations >6.5 mg/L. This subdividing of the sample focused the final discriminant analysis on a residual group (subgroup 3) of cases with PG <0.5 mg/L and sphingomyelin ≤6.5 mg/L.

The decile frequencies revealed that for subgroup 3, L/S now showed a misclassification rate of 18.3%, while lecithin showed a rate of 22.6% (Tables 4, 5). The stepwise discriminant analysis, however, indicated that lecithin concentration, when combined with L/S, no longer made a significant contribution to the discriminant. It appeared that the subgrouping procedure, by eliminating the cases with high PG (and therefore high lecithin concentrations) in step one, had already taken into consideration this independent discriminant information. For purposes of comparison, we have shown the values calculated from the linear function derived from the step-wise analysis on subgroup 3:

$$MX3 = (L/S \cdot 0.712) + (Lec \cdot 0.0189) - 2.262$$  \hspace{1cm} (3)

### Table 2. Discriminant Power of Tests for Total Sample

<table>
<thead>
<tr>
<th>Test</th>
<th>No.</th>
<th>% of total</th>
<th>Mean RDS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>L/S ratio</td>
<td>50</td>
<td>14.4%</td>
<td>3.87</td>
<td>1.87</td>
<td>59.8</td>
</tr>
<tr>
<td>Lecithin</td>
<td>64</td>
<td>18.4%</td>
<td>34.9</td>
<td>14.3</td>
<td>47.3</td>
</tr>
<tr>
<td>Phos. glycerol*</td>
<td>92</td>
<td>26.6%</td>
<td>2.65</td>
<td>1.92</td>
<td>0.08</td>
</tr>
<tr>
<td>Sphingomyelin*</td>
<td>100</td>
<td>28.9%</td>
<td>10.4</td>
<td>4.19</td>
<td>1.54</td>
</tr>
<tr>
<td>Disc. func. 1</td>
<td>48</td>
<td>13.8%</td>
<td>2.31</td>
<td>1.71</td>
<td>71.7</td>
</tr>
<tr>
<td>Disc. func. 2</td>
<td>52</td>
<td>15.0%</td>
<td>1.95</td>
<td>-0.97</td>
<td>65.2</td>
</tr>
</tbody>
</table>

*Units for these in columns 4 and 5 are mg/L.

### Table 3. Subdividing Total Sample by Different Tests

<table>
<thead>
<tr>
<th>Group</th>
<th>Definition</th>
<th>No. cases</th>
<th>No. with RDS</th>
<th>Percentage with RDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PG ≥0.5 mg/L</td>
<td>131</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>S &gt;6.5 mg/L</td>
<td>29</td>
<td>13</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Residual</td>
<td>186</td>
<td>42</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>346</td>
<td>58</td>
<td>17</td>
</tr>
</tbody>
</table>
Both this and the previously derived function, MX1, also showed a misclassification rate of 18.3%. We further confirmed that lecithin concentration was not a significant factor in the discriminant for subgroup 3 by testing, stepwise, the combination of L/S, followed by MX1. The MX1 function, combining L/S and LEC, did not add significantly to the discriminant achieved with L/S alone.

In subgroup 3, S again acted as a negative control, showing no discriminatory power, owing to a 36.6% misclassification rate. The expected misclassification rate for an ineffective test was calculated to be 35%. However, subgroup 2 showed a RDS rate of 64% in cases with L/S of <2.0, and a rate of 26% in cases with an L/S of 2.0 or greater. Thus, high sphingomyelin concentrations in the amniotic fluid of these patients without detectable concentrations of PG indicated an increased risk of RDS in the infant. The high rate of RDS in this subgroup resulted from the relatively low average gestational age, approximately 33 weeks, vs 35 weeks for cases in subgroups 1 and 3. In most of these cases, complicating factors had necessitated early delivery. Although the number of cases in this group was too small for any further subclassification, some diagnoses appeared with unexpected frequency: infection (six cases), congenital defects (four cases), isoimmunization (three cases), abruptio placenta or placenta previa (three cases), and history of previous fetal loss (three cases). Apparently, these pathological processes that necessitated early delivery frequently contributed contaminating phospholipids to the amniotic fluid.

Discussion

Although many tests on amniotic fluid distinguish with various degrees of success between RDS and pulmonary maturity in the fetus, it has been difficult to obtain valid evaluations measuring the comparative effectiveness of these tests. This is true for many reasons. Accumulating data on a large group of patients requires a long time. The incidence of RDS is usually low. Definitions of RDS, as distinguished from TTN, vary from study to study, making intercomparisons of studies almost impossible. In addition, tests are frequently done without adequate standardization, creating even greater difficulty in interlaboratory comparison (13)—a problem further compounded when workers develop their own modifications of a particular procedure, yielding highly diverse results (14, 15). Finally, the cutoff value for various tests, the value used in distinguishing disease from non-disease, is often arrived at arbitrarily by each investigator, most often reflecting the balance between the number of false positives vs the number of false negatives that the investigator is willing to accept. By shifting this decision point separating disease from non-disease, one can greatly increase either sensitivity or specificity, but the gain in one is always achieved at the expense of the other. Thus, comparisons based on reported sensitivity or specificity of a test are frequently meaningless.

Table 5. Discriminant Power of Tests on Group 3

<table>
<thead>
<tr>
<th>Test</th>
<th>Misclassified cases</th>
<th>Mean result</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>L/S ratio</td>
<td>No. % of total</td>
<td>Mature</td>
<td>RDS</td>
<td>F</td>
</tr>
<tr>
<td>Lecithin*</td>
<td>42 %</td>
<td>22.6</td>
<td>21.3</td>
<td>8.76</td>
</tr>
<tr>
<td>Sphingomyelin*</td>
<td>68 %</td>
<td>36.6</td>
<td>3.33</td>
<td>3.60</td>
</tr>
<tr>
<td>Disc. Func. 1</td>
<td>34 %</td>
<td>18.3</td>
<td>2.44</td>
<td>1.49</td>
</tr>
<tr>
<td>Disc. Func. 3</td>
<td>34 %</td>
<td>18.3</td>
<td>2.54</td>
<td>-8.7</td>
</tr>
</tbody>
</table>

*Units in columns 4 and 5 are mg/L.
ing between RDS and pulmonary maturity is difficult to confirm in other laboratories.

In this study we have tried to deal with some of the difficulties noted above. All assays were standardized against reference phospholipids and were checked with control materials on every chromatographic plate. We measured both the spot reflectance ratio and the actual concentrations of lecithin and sphingomyelin. It is apparent that reference standards should be routinely used in this assay, when L/S results for a specimen submitted to different laboratories may range from 1.3 to 10.0 (15). Some workers have measured the actual concentrations of phospholipids in the amniotic fluid vs reference standards (5, 28–32), but the ratio of spot densities of lecithin and sphingomyelin (in arbitrary units) is still the routine approach used in most clinical laboratories. Comparing clinical results with different unreferences methodologies that vary from laboratory to laboratory is unlikely to be widely meaningful. Our data illustrate this point (Table 4). We observed a rate of RDS of about 17% for L/S 2.0–2.8, contrasting with papers that report RDS only infrequently for L/S in this range. Two factors—one technical, the other clinical—appeared to account for these different observations. We followed the original method of Gluck et al. (2), spraying the plates with 9 mol/L sulfuric acid before charring. We have found this step to be essential if spot density is to remain proportional to the amount of phospholipid (7, 9). Without such spraying, equal amounts of amniotic fluid lecithin, sphingomyelin, and PG—as well as standards of dipalmitoyllecithin and sphingomyelin—all give different spot densities, making accurate quantification impossible. Also, RDS, unlike hyaline membrane disease, is a clinical syndrome that is not precisely defined; our definition included cases that were treated with low amounts of supplemental oxygen for 48 to 72 h, cases that some workers, using more severe criteria, would have excluded from the analysis. This combination of technical and clinical factors placed more cases of RDS in the L/S range of 2.0–2.8 in our study. It is quite likely that others, using only slightly less sensitive assay methods and clinical definitions, would find almost no cases of RDS in this range.

By performing the different assays on the same specimens, we were also able to use a common decision point when we compared the effectiveness of each test and group of tests. These analyses indicated that L/S is a more effective test than lecithin concentration. PG, correlating strongly with high lecithin concentrations, apparently contributed nothing to the discriminant under such circumstances. The combination of lecithin concentration and L/S was slightly more effective than either one alone, indicating that, at least in some cases, lecithin concentration and L/S are providing independent diagnostic information. Although the combination, L/S and lecithin concentration, appeared to offer a small but statistically significant advantage, the failure of the combination to decrease the number of misclassified cases suggested that this improvement in diagnosis is of minimal clinical significance.

The phosphatidylylycerol and sphingomyelin assays were effective subgrouping procedures, however, segregating groups with high and low probability of RDS. Because all these phospholipids can be quantified by running standards in three channels of the chromatographic plate (7), it would appear that this simple addition to the Gluck procedure would be justified on the basis of the increment in diagnostic information. Of particular importance, standardization permitted detection of specimens contaminated with non-surfactant phospholipids. Although contamination was frequently obvious from the color of the amniotic fluid specimen, our observations suggested that about half of the specimens with high sphingomyelin concentrations were not off-color, and, equally important, some off-color specimens did not contain significantly increased concentrations of sphingomyelin. We consider sphingomyelin to be the more reliable indicator of specimen contamination, at least in terms of interpreting the test results and predicting RDS.

Finally, these subgrouped data suggested a simple stepwise strategy for optimizing the evaluation of pulmonary function in the fetus:

1. measure the concentration of phosphatidylglycerol
2. measure the concentration of sphingomyelin
3. evaluate the L/S ratio

A PG of 0.5 g/L or more (subgroup 1) indicates about a 1 to 2% probability of RDS. In this regard, we have found an immunoassay technique for quantifying PG a useful first procedure (32). If PG is not present—i.e., less than 0.5 mg/L—then sphingomyelin is evaluated to determine if the amniotic fluid contains contaminating phospholipids (subgroup 2). In this group, the L/S ratio must be cautiously interpreted. If it is <2.0, the probability of RDS is about 65%; if >2.0, RDS is found in about 25% of the cases. Finally, in the residual cases with no PG and normal sphingomyelin (subgroup 3), the probability of RDS can be estimated from the L/S ratio: about 70% for values ≤1.7, gradually decreasing to about 1% when the ratio reaches 2.8. However, we emphasize that, although this general approach would optimize discriminating between RDS and normal pulmonary function in all laboratories, the exact ranges for L/S would depend upon the method used for making the spots visible and the definition used for RDS.

References

CLINICAL CHEMISTRY, Vol. 33, No. 5, 1987