Answers to Evaluation Questions

1. a. An example of a completed table, based on data shown in Figures 24.3 and 24.5, is shown below.

<table>
<thead>
<tr>
<th>Session</th>
<th>Intercept (constant bias)</th>
<th>Slope (systematic bias)</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>–20.23</td>
<td>1.31</td>
<td>87.4%</td>
</tr>
<tr>
<td>Posttest</td>
<td>2.54</td>
<td>0.88</td>
<td>93.2%</td>
</tr>
<tr>
<td>Improvement (yes/no)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

In this example, we can state that training with DiseasePro did improve the accuracy of the posttraining assessments. Comparing the two sessions, we see that in the posttest session the value of the slope was closer to 1.0, the \( y \)-intercept was closer to 0, and the coefficient of determination \( (R^2) \) was higher.

b. The regression equations shown in Figure 24.5 have intercepts that are different from 0 and slopes that are different from 1, so both constant bias and systematic bias are present in the estimates. In pretraining assessments, constant bias was present in the estimates (intercept = –20.23%) and the rater’s response to increasing levels of disease severity showed systematic bias (slope = 1.31%). After training, the rater’s assessments indicated a much lower amount of constant bias (intercept = 2.54%) and reduced systematic bias (slope = 0.88%).

c. The example pretest error graph in Figure 24.5A shows a definite, nonlinear pattern in the error terms (estimated severity minus actual severity). As actual disease severity increased between 0 and 30%, estimate errors increased, and as actual severity increased from 30 to 60%, estimate errors decreased. This rater consistently underestimated disease severity.
severity throughout the test, but estimates given for images with relatively low (0–10%) or high (50–60%) levels of severity were slightly closer to the true values. After training (Fig. 24.5B), this rater’s estimates were within 10% of the true values and the error pattern was nearly random.

d. On the basis of the pretest error graph, this rater was least accurate when actual disease severity was close to 30%.

2. The student should compare estimate error graphs, point out differences or similarities in estimate errors, and describe how errors changed as actual severity increased.

a. The student should contrast and compare his or her pretest regression results (slope, intercept, and $R^2$) with those of another student and discuss the results in terms of systematic bias and constant bias. For example, assume that the two data plots shown in Figure 24.1 are the pretest results for two raters. Neither displays constant bias (both intercepts are very close to 0), but both display systematic bias. Rater A is more likely to overestimate as “true” severity increases (slope = 1.2%), and rater B is more likely to underestimate as severity increases (slope = 0.75%). A higher $R^2$ value indicates higher precision. Rater A’s estimates ($R^2 = 97.88\%$) are slightly more precise than those of rater B ($R^2 = 96.63\%$). Because the slope and $R^2$ values from rater A’s estimates are closer to the standards for perfect accuracy (slope = 1, $R^2 = 100\%$), we can state that rater A’s estimates are more accurate than those of rater B.

b. In Figure 24.1, the two sets of estimates do not show good agreement; the difference between estimates increases as actual disease severity increases. The regression lines indicate that rater A’s estimates are higher than the actual severity level of 25%, whereas rater B’s estimates are low when actual severity was 25%.

c. In the example shown in Figure 24.1, assessments made by rater A would trigger use of the management practice before it was really needed. Rater B underestimates actual severity, so assessments made by rater B would not trigger the action threshold at the correct time and action would be delayed.

d. Refer to your posttest results and those of a second person, and discuss how the accuracy of management recommendations was affected by assessment training.

(Optional)

3. An example of a combined regression graph is shown in Figure 24.7. Compare 95% confidence intervals for regression parameters, report your comparison, and provide copies of the statistical analysis output (not shown). State how you compared the difference between parameters (e.g., were interpretations based on visual inspection of the 95% confidence intervals or on statistical tests of the difference?).

a. Report the SEEy statistics for your pretest and posttest sessions, and interpret these values in terms of estimate accuracy. A lower SEEy value indicates that estimates were closer to actual disease severity levels.

Answers are not provided for the optional exercises in 3.b and 3.c. Be prepared to provide copies of your test results, error graphs, regression graphs, and statistical analyses, and interpret how your rating accuracy changes when you are assessing a different lesion size or new disease.