

The Shewhart p Chart for Comparisons

by Marilyn K. Hart and Robert F. Hart

An important question regarding test result validity is, Given a positive test result, what is the probability that this test result is wrong? This question is answered by the *positive predictive error rate*, the number of false positive test results divided by the total number of positive test results.

The data used here were first included in a case study in *Statistical Process Control for Health Care* by Marilyn K. Hart and Robert F. Hart. Table 1 presents data for evaluating the positive predictive ultrasound error rate for seven radiologists in their ultrasound testing for acute appendicitis. The first three radiologists are radiology specialists; the last four are nonspecialists. Table 2 summarizes these results for the two groups.

Table 1. Acute Appendicitis: Positive Predictive Errors by Radiologist

Radiologist	False Positives	Total Positives
1	1	8
2	4	12
3	2	8
4	3	5
5	5	7
6	2	6
7	2	4

Table 2. Acute Appendicitis: Positive Predictive Errors by Radiologist Group

Radiologist Group	False Positives	Total Positives
Specialists	7	28
Non-specialists	12	22

One might make a p chart with seven subgroups to compare the radiologists and/or with two subgroups to compare the groups. It is common for time-ordered control charts to have 25 subgroups, in which case the common 3-sigma limits are appropriate. However, for fewer subgroups, 3-sigma limits are too wide to be effective. Recommended values of T for T-sigma limits are given in Table 3.

Table 3. Process Evaluation for Special-cause Variation:
Recommended Values of T for T-sigma Limits*

# of subgroups	T
2	1.5
3-4	2.0
5-9	2.5
10-34	3.0
35-199	3.5
200-1500	4.0

* The tabular values of T may be used for the usual case of “no standard given” with all attribute and variables charts.

A common problem with p charts is that the subgroup sizes are too small to give valid results. The required minimum subgroup size is a function of pBar. For a point to be accepted as valid, a minimum subgroup size of $1/p\text{Bar}$ is needed and for a point above the upper control limit to be accepted as valid, a minimum subgroup size of $4/p\text{Bar}$ is needed. Here the pBar for the error rate is 0.38 (whether one subgroups the data by radiologist or by radiologist group).

The p chart subgrouped by radiologist (not shown here) shows that the 7 points all fall below the 2.5-sigma upper control limit with the subgroup sizes all adequate (i.e., exceeding the minimum requirement of $1/0.38 = 2.63$ rounded up to three). With only this analysis one would conclude that special-cause variation between radiologists was not demonstrated. This does not mean that no special-cause variation existed—only that a more powerful method of subgrouping was needed to flush out the lack of statistical control.

Figure 1 is a p chart with 1.5-sigma limits profiling the two radiologist groups on their positive predictive error rate. This chart illuminates the superior performance of the specialists, which should be no surprise. It should be noted that the out-of-control condition is to be accepted as valid since the minimum required subgroup size of $4/0.38 = 10.53$ is met.

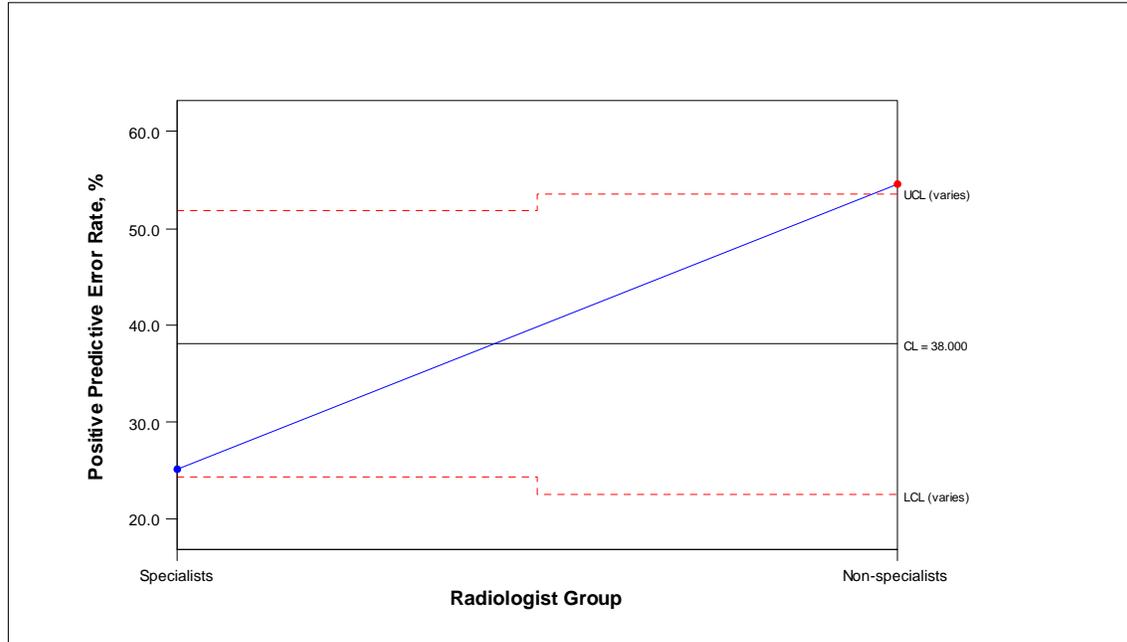


Figure 1. p Chart Profiling Radiologist Group on Positive Predictive Error Rate, 1.5-sigma Limits

The p chart on individual radiologists had insufficient power to detect any special-cause variation. By subgrouping the seven radiologists into specialists and nonspecialists, the larger subgroup sizes provided the increased power needed to detect the special-cause variation.

Should only specialists perform the ultrasound tests?

References

Hart, Marilyn K., and Robert F. Hart. *Statistical Process Control for Health Care*. Pacific Grove, California: Duxbury, 2002.

Copyright © 2006 by Marilyn K. Hart and Robert F. Hart.

Stay informed on quality efforts in healthcare.
Sign up to receive ASQ's *Healthcare Update Newsletter*:
http://www.asq.org/healthcare/update_info.html

About the Authors

Robert F. Hart, Ph.D., and Marilyn K. Hart, Ph.D., have more than 20 years of experience in statistical process control and have taught more than 300 seminars for healthcare and for industry. They have worked with healthcare organizations for more than 12 years, teaching seminars, analyzing data, and performing collaborative research with various organizations, including JCAHO.

Marilyn K. Hart has an academic background in mathematics and operations research with a doctorate in management sciences from the Illinois Institute of Technology. A student of Dr. W. Edwards Deming, she has taught in higher education, industry, and healthcare; has consulted in manufacturing and healthcare; and is a professor in the College of Business Administration, University of Wisconsin Oshkosh.

Robert F. Hart has a doctorate from Northwestern University in engineering with a subsequent master's degree in operations research. A student of Dr. W. Edwards Deming, he worked with General Motors for 20 years, has been a consultant since 1980, and is an adjunct professor in the College of Business Administration, University of Wisconsin Oshkosh. Dr. Robert Hart has completed the Intermountain Health Care Advanced Training Program in Health Care Delivery Improvement.