

Guidance 089 (QRM) Application to Establishment of Weighing Device Performance Testing Intervals

The probability of the weighing device encountering drift during the proposed period should be based on the operating history of the device. This review should include the frequency of unscheduled maintenance over the previous two year period and the device's performance history from, minimally, the last five calibrations. If the idealized running history, respective of maintenance and calibration, is not available for a specific weighing device, the history of a like model instrument deployed under similar conditions may be used for reference.

The most critical measure is of the performance testing history. This assessment should be based upon a statistically significant number of consecutive measures; a minimum of 30 consecutive measurements (representing the current verification testing frequency, e.g. daily) is recommended to minimize the influence of statistical error. All data points for each weight employed during the performance verification test should fall within $\pm 3 \sigma$ of the overall mean for each weight. Any points falling outside the $\pm 3 \sigma$ range should be investigated.

Business – although decreasing the test frequency may save time and money in normal practice, it must be recognized that the inverse would be true when a failure is observed during less frequent (e.g. weekly) testing. A greater number of weighing operations may be termed “suspect” and subsequently require repeat. This would represent greater impact to cycle times, costs of goods, and man hours required for investigations and repeat operations. Therefore as the frequency of use for a weighing device increases, so does the potential impact of a performance failure.

For each of the above stated risks related to a change in frequency of performance verification testing the individual risk components must be assessed. As identified previously each potential risk has an associated probability and a severity. Table 1 below represents the suggested scale for use in the assessment; it reflects the differentiations in harm resulting in each area of risk relative to the scenario. During the assessment the risks will be scored based on the worst applicable harm associated with a given risk.