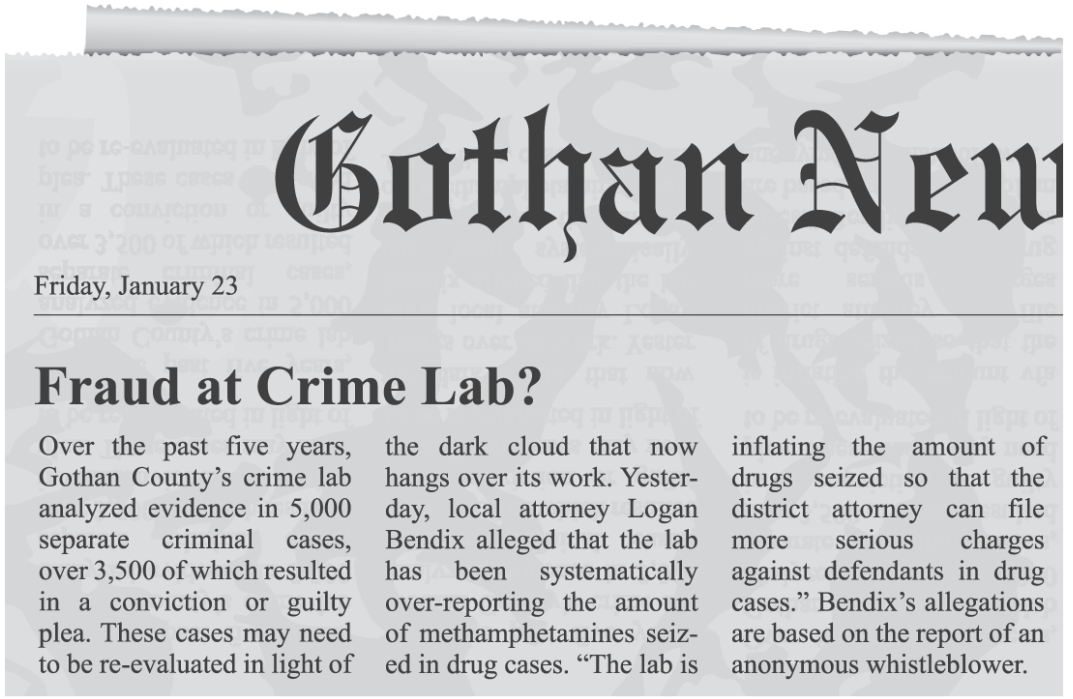
Quality Assurance and Calibration Methods

Chapter 5

Case Study: Misconduct at the Gothan Crime Lab?

How to determine if data was falsified?

- Senior technician is accused of falsifying methamphetamine assays
- Accuser is a visiting scientist and may/may not have ulterior motives
- How do you determine (1) if mistakes are made and (2) if mistakes are intentional?



Gothan New

Friday, January 23

Fraud at Crime Lab?

Over the past five years, Gothan County's crime lab analyzed evidence in 5,000 separate criminal cases, over 3,500 of which resulted in a conviction or guilty plea. These cases may need to be re-evaluated in light of the dark cloud that now hangs over its work. Yesterday, local attorney Logan Bendix alleged that the lab has been systematically over-reporting the amount of methamphetamines seized in drug cases. "The lab is inflating the amount of drugs seized so that the district attorney can file more serious charges against defendants in drug cases." Bendix's allegations are based on the report of an anonymous whistleblower.

S. M. "Contakes, Misconduct at the Lab?", *J. Chem. Ed.* **2016**, *93*, 314 (supplementary material file 1) https://www.fodey.com/generators/newspaper/snippet. asp. Reprinted with permission © 2016 American Chemical Society

- Conclusion could force re-examination of 3 000+ drug cases

Section 5-1
Basics of Quality Assurance

Quality Assurance

Quality assurance is what we do to get the right answer.

- Answer should have sufficient accuracy and precision to support subsequent decisions
- No point in obtaining more accurate/precise answer than necessary

Data quality standards:

- Get the right data.
- Get the data right.
- Keep the data right.

(Nancy W. Wentworth, U.S. EPA)

Table 5-1 Quality assurance process

Question	Actions
Use Objectives Why do you want the data and results, and how will you use the results?	• Write use objectives
Specifications How good do the numbers have to be?	• Write specifications • Pick methods to meet specifications • Consider sampling, precision, accuracy, selectivity, sensitivity, detection limit, robustness, and rate of false results • Employ blanks, fortification, calibration checks, quality control samples, and control charts to monitor performance • Write and follow standard operating procedures
Assessment Were the specifications achieved? Is the method fit for purpose?	• Compare data and results with specifications • Document procedures and keep records suitable to meet use objectives • Verify that use objectives were met

Use Objectives

The goal of quality assurance is making sure that results meet the customer's needs. Write clear, concise **use objectives** and keep in mind:

Raw data: measurements

Treated data: concentrations derived from raw data by use of calibration methods

Results: quantities reported after statistical analysis of treated data

Specifications

Specifications: state how good the numbers need to be and what precautions are required in the analytical procedure

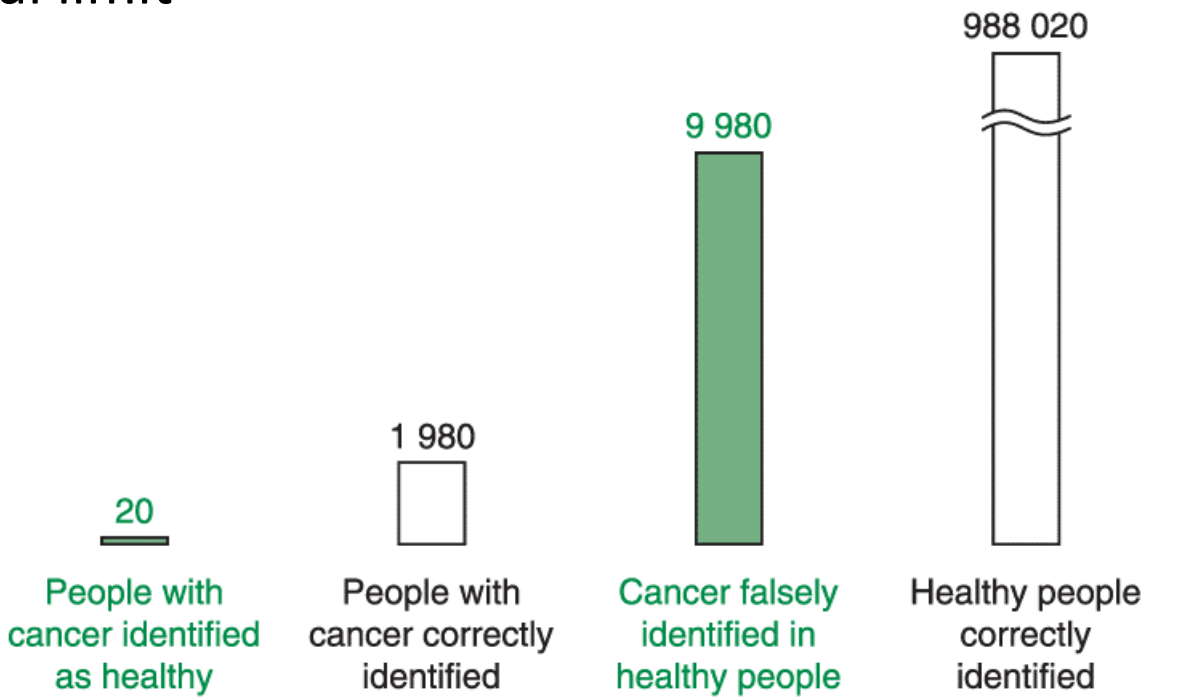
- How shall samples be taken?
- How many are needed?
- Are special precautions needed to avoid degradation?
- What are practical restraints (cost, time, material)?
- What level of accuracy will satisfy the use objectives?
- What rate of false negatives or positives is acceptable?

Box 5-1 Medical Implication of False Positive Results

False positive: concentration exceeds the legal limit when, in fact, the concentration is below the limit

False negative: concentration is below the legal limit when it is actually above the limit

Even well-executed procedures produce some false conclusions due to random error.



Harris/Lucy, *Quantitative Chemical Analysis*, 10e, © 2020 W. H. Freeman and Company

When Choosing a Method Consider

Selectivity: ability to distinguish analyte from other species in sample (avoid interference)

Sensitivity: ability to respond reliably and measurably to change in analyte concentration (slope of the calibration curve)

Blanks: sample not intended to contain analyte

- Account for interference by other species in the sample
- Account for trace amount of analyte found in reagents
- Frequent measures of blanks detect whether analyte from previous sample is carried into subsequent analysis by vessels or instruments

Types of Blanks

Method blank

- All components except analyte
- Taken through all steps of the analytical procedure
- Subtract the response of the method blank from the response of sample before calculating the quantity of analyte

Reagent blank

- Similar to a method blank, but it has not been subjected to all sample preparation procedures

Field blank

- Indicates if analyte is inadvertently picked up by exposure to field conditions

Spike Recovery

Sometimes response to analyte can be decreased or increased by something else in the sample

Matrix: everything in the unknown, other than the analyte

Spike (or fortification): a known quantity of analyte added to a sample

- Tests whether the response to the spike is the same as expected based on known calibration curve
- Response is different from expectations indicates problems (contamination or loss)

Example: Spike Recovery (1 of 3)

Let *C* stand for concentration. One definition of spike recovery is

$$\%\ \text{recovery} = \frac{C_{\text{spiked sample}} - C_{\text{unspiked sample}}}{C_{\text{added}}} \times 100$$

An unknown was found to contain 10.0 µg of analyte per liter. A spike of 5.0 µg/L was added to a replicate portion of unknown. Analysis of the spiked sample gave a concentration of 14.6 µg/L. Find the percent recovery of the spike.

Example: Spike Recovery (2 of 3)

Solution: The percent of the spike found by analysis is

$$\% \text{ recovery} = \frac{14.6\ \mu g/L - 10.0\ \mu g/L}{5.0\ \mu g/L} \times 100 = 92\%$$

If the acceptable recovery is specified to be in the range of 96 to 104%, then 92% is unacceptable. Something in your method or techniques needs improvement.

Example: Spike Recovery (3 of 3)

Test Yourself: Find percent recovery if the fortified sample gave a concentration of 15.3 µg/L.

Dealing with Large Numbers of Samples

Perform periodic calibration checks.

Method drift can be due to changes such as room temperature or spoilage of reagents/standards.

Calibration check solutions should be different from ones used to prepare the original calibration curve.

Performance test samples (quality control samples or blind samples) help to eliminate bias introduced by an analyst who knows the concentration of the calibration check sample.

Standard operating procedures (SOP) outline steps to be taken for the procedure. Serve as control experiments to detect problems in the lab.

Assessment

Assessment is the process of

(1) Collecting data to show that analytical procedures are operating within specific limits

(2) Verifying that final results meet use objectives

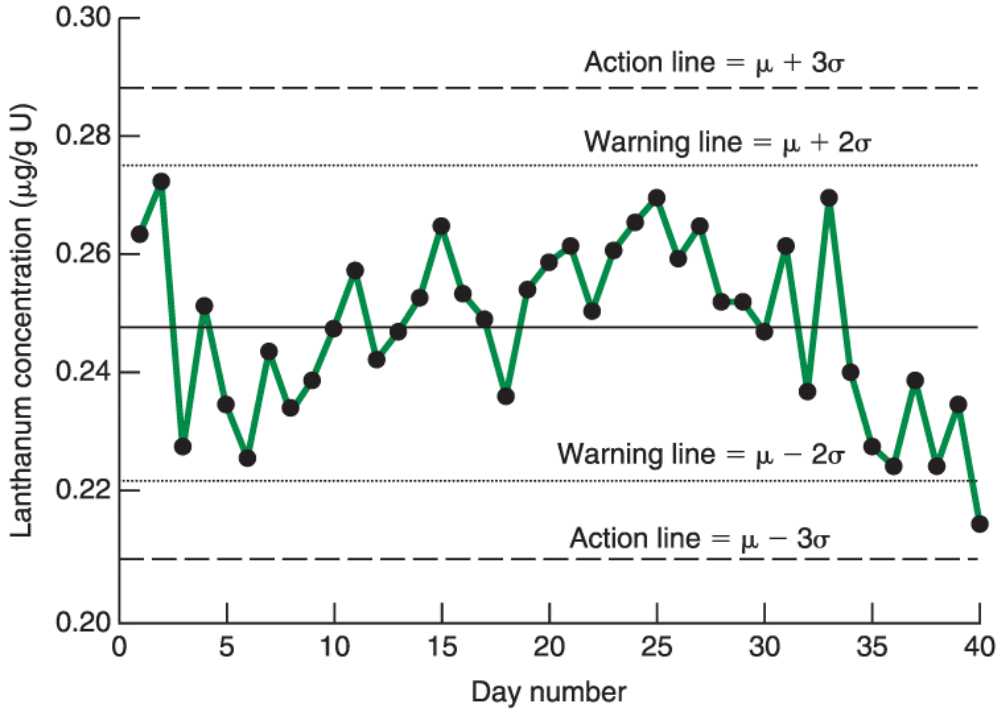
Documentation is critical. Standard protocols provide direction for what must be documented and how documentation is to be done.

Control charts can be used to monitor performance.

If final results meet the use objectives, the method is **fit for purpose**.

Box 5-2 Control Charts

Laboratory measuring lanthanum (La) in uranium as part of international nuclear non-proliferation monitoring



Harris/Lucy, *Quantitative Chemical Analysis*, 10e, © 2020 W. H. Freeman and Company

- Control chart: visual representation of confidence intervals for a Gaussian distribution
- Warns when a property strays dangerously far from an intended *target value*
- For a Gaussian distribution, 95.5% of all observations are within $\pm 2\sigma$; 99.7% within $\pm 3\sigma$
- $\pm 2\sigma$ = **warning lines** and $\pm 3\sigma$ = **action lines**

Table 5-2 Agencies that develop standard methods (1 of 2)

Government agencies such as the U.S. Environmental Protection Agency (EPA) set requirements for quality assurance and publish **standard methods** for certified analysis.

American Public Health Association (APHA)

- 400 methods for analysis of water, water supplies, and wastewater

AOAC International (formerly Association of Official Analytical Chemists)

- 3 000 standardized chemistry and microbiological methods to ensure safety of food, beverages, dietary supplements and similar products, and purity of their ingredients

ASTM International (formerly American Society for Testing and Methodology)

- 5 400 test methods for 90 industrial sectors, including oil and gas, mining, pulp and paper, industrial chemicals, agriculture, and energy

National Institute for Occupational Safety and Health (NIOSH)

- Methods for industrial hygiene monitoring

Table 5-2 Agencies that develop standard methods (2 of 2)

Occupational Safety and Health Administration (OSHA)

- Methods for sampling and analysis of contaminants in workplace air, on workplace surfaces, and in the blood and urine of workers who are occupationally exposed

U.S. Environmental Protection Agency (EPA)

- 1 600 methods for drinking water, air pollution, water pollution, hazardous waste, pesticides, and radiochemistry

U.S. Food and Drug Administration (FDA)

- Methods for allergens, additives, supplements, pesticides/herbicides, drug residues, toxic elements, bacteria, and microbiologics in food and cosmetics
- Regulations and guidelines for the validation of analytical methods and procedures for drugs and biologics

U.S. Pharmacopeia (USP)

- Standards for medicines, food ingredients, dietary supplement products, and ingredients
- Other similar agencies are the British Pharmacopeia and European Pharmacopeia

Section 5-2
Method Validation

Method Validation

Method validation: the process of proving that an analytical method is acceptable for intended purpose

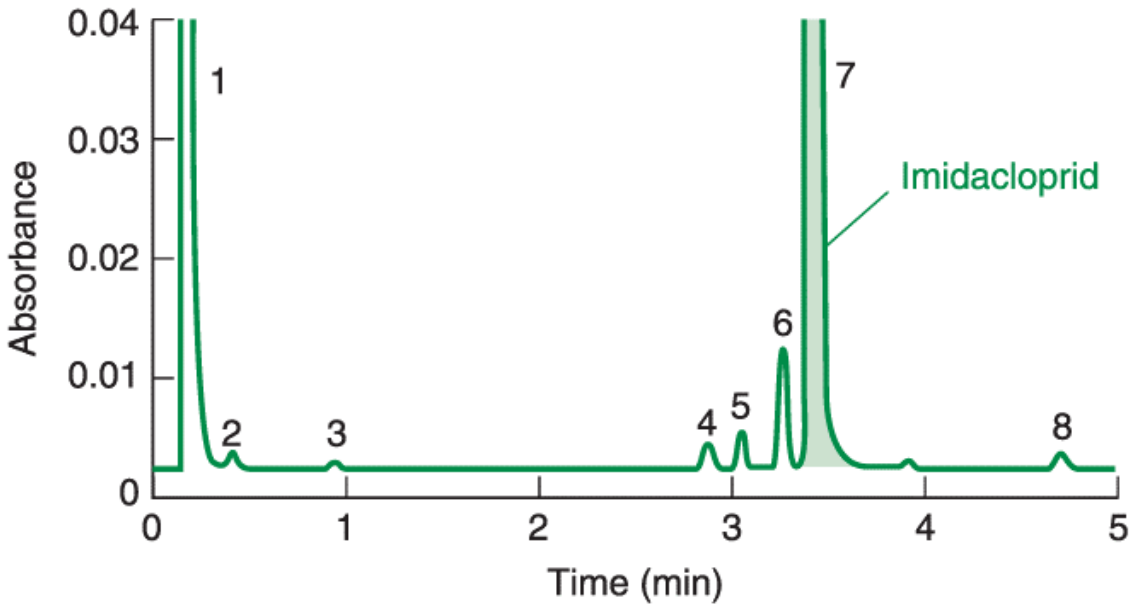
- Method selectivity
- Linearity
- Accuracy
- Precision
- Range
- Limit of detection
- Limit of quantitation
- Robustness

Selectivity

Selectivity: extent to which an analytical method can distinguish analyte from everything else in the sample

- Method is **specific** if it selects for only one analyte (no interferences).
- Chromatogram separates drug imidacloprid (peak 7) from other potential impurities/degradation peaks.

Figure 5-1



Harris/Lucy, *Quantitative Chemical Analysis*, 10e, © 2020 W. H. Freeman and Company

Linearity

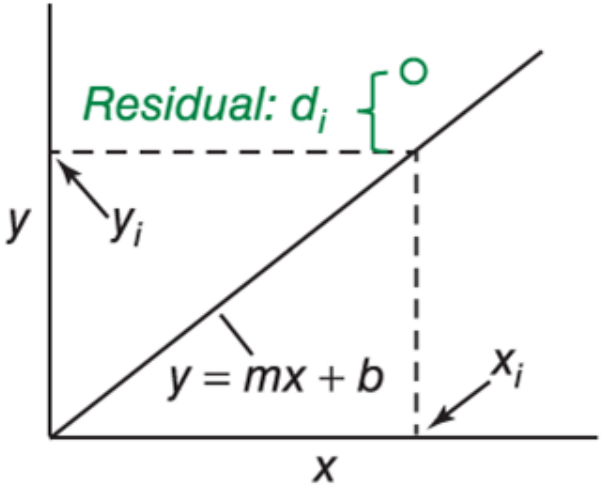
Linearity: measures how well a calibration curve follows straight line

➔ response is proportional to quantity of analyte

A common (but superficial) measure of linearity is the **square of the correlation coefficient, R^2**.

$$R^2 = \frac{\left[\sum(x_i - \bar{x})(y_i - \bar{y})\right]^2}{\sum(x_i - \bar{x})^2 \sum(y_i - y)^2}$$

- A value of $R^2 > 0.995$–0.999 is a good fit for many purposes.
- Plotting a graph shows the true nature of the calibration.

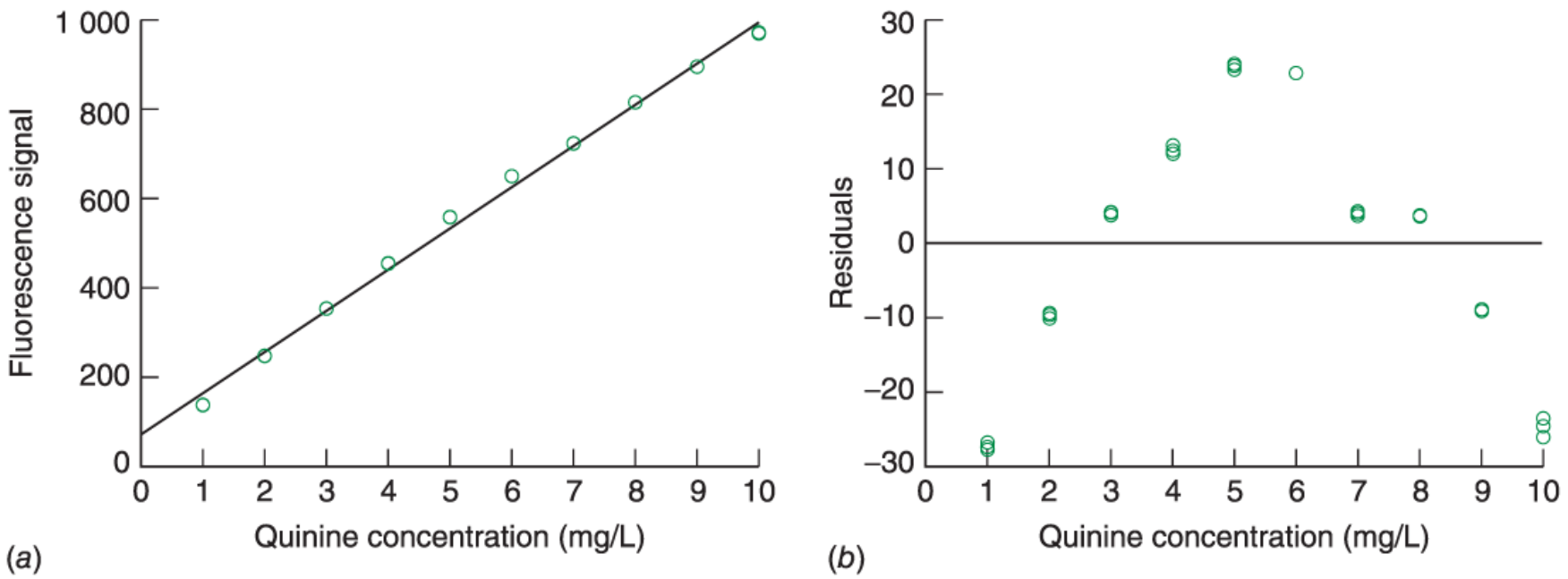


Harris/Lucy, *Quantitative Chemical Analysis*, 10e, © 2020 W. H. Freeman and Company

Residual Plots

Residual plots emphasize the difference between calibration data and the least-squares line.

Figure 5-2



Harris/Lucy, *Quantitative Chemical Analysis*, 10e, © 2020 W. H. Freeman and Company

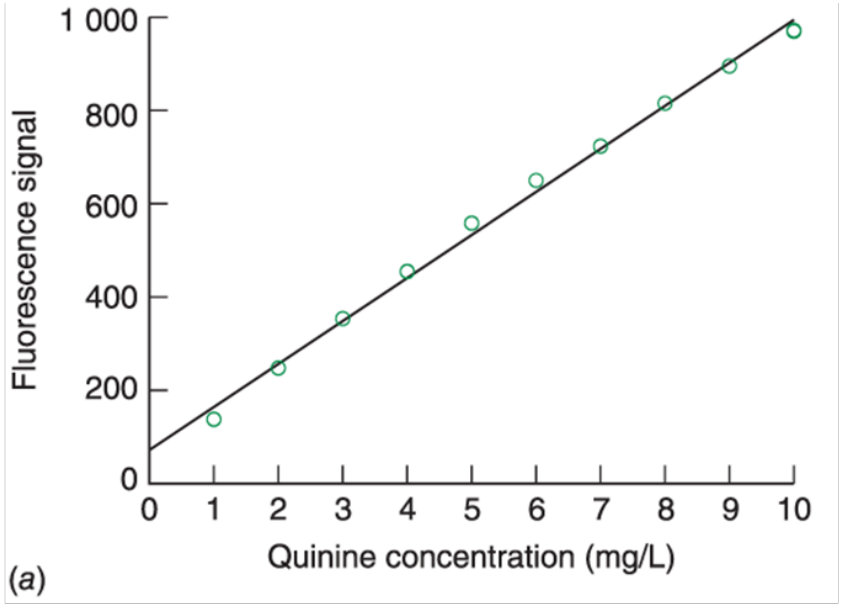
Residual

$$d_i = y_i - y$$
$$= y_i - (mx_i + b)$$

Example: Linearity (1 of 3)

Samples from 2 to 3 mg/L quinine are to be analyzed using fluorescence spectroscopy. Standards from 1 to 10 mg/L were measured in triplicate to yield the calibration curve in Figure 5-2a. Least-squares yields $y = 92.4x + 72.9$ with $R^2 = 0.996$. Is the calibration linear?

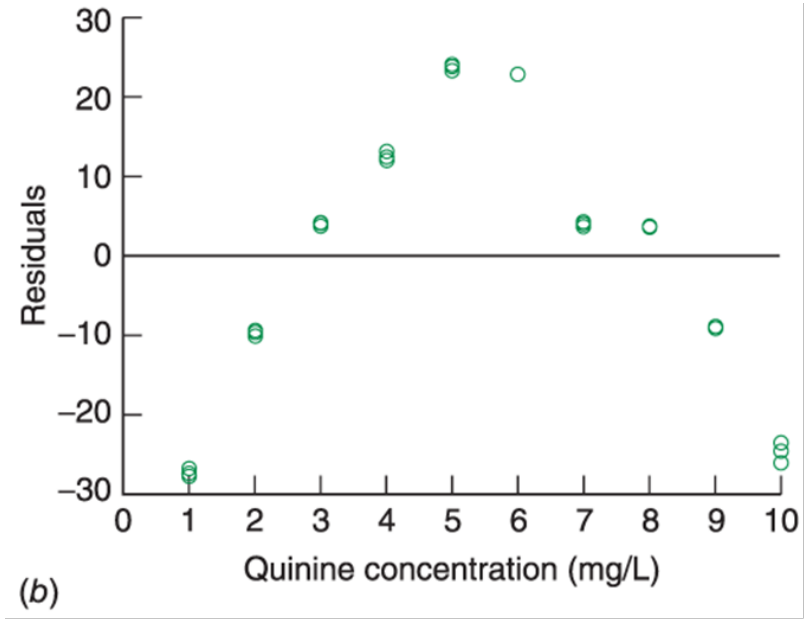
Figure 5-2a



Example: Linearity (2 of 3)

Solution: R^2 is greater than 0.995, but R^2 alone should not be used to judge linearity. The intercept is 21% of the 3-mg/L standard, which is greater than the ≤2% criterion. Inspection of the calibration plot in Figure 5-2a reveals that low standards are below the line, middle standards are above the line, and high standards are below the line. The residual plot in Figure 5-2b clearly indicates curvature.

Figure 5-2b



Example: Linearity (3 of 3)

Solution: Fitting the calibration data to a quadratic function yields $y = -2.25x^2 + 117.1x + 23.4$, with $R^2 = 0.999\ 8$. Residuals for the quadratic fit are randomly scattered about 0 and reduced to one-third of those in Figure 5-2b. The intercept is still >2% of the 3-mg/L standard.

Reviewing the procedure revealed that fluorescence was zeroed using distilled water rather than the 0.05 M sulfuric acid used to prepare standards. Sulfuric acid may have given the fluorescence responsible for the positive intercept. In this case, deviation of the intercept from 0 cannot be used to assess the fit. Based on these considerations, the quadratic fit is appropriate.

Accuracy: “Nearness to the Truth”

Accuracy can be demonstrated by:

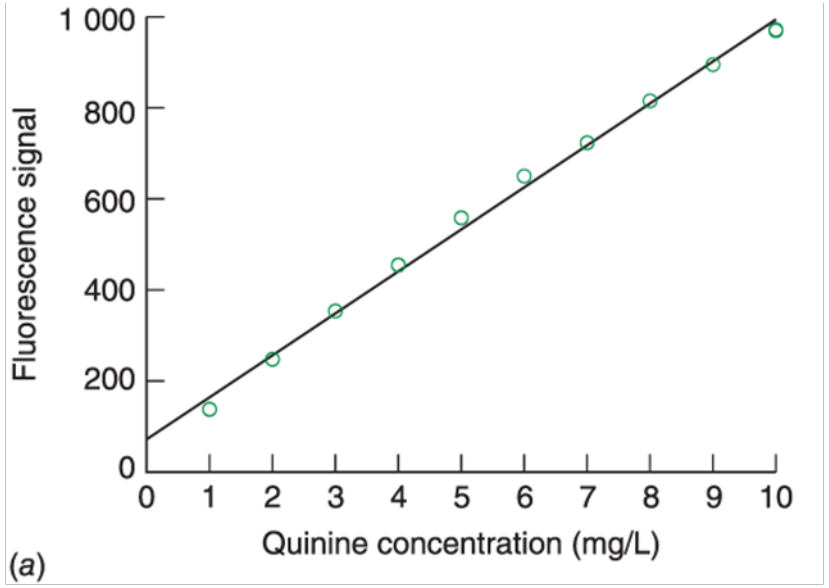
- Analyzing a certified reference material in a matrix similar to that of the unknown
- Comparing results from two or more different analytical methods
- Analyzing a blank sample *spiked* with known addition of analyte
- *Standard additions* of analyte to the unknowns

Example: Testing Accuracy (1 of 3)

A specification for determining ~3 mg/L quinine by fluorescence in the preceding example is a spike recovery of 100 ± 2%. Using the quadratic fit for the 1- to 10-mg/L calibration, we can estimate that blank solutions spiked to 1.50 and 4.50 mg/L quinine would yield signals of 194.0 and 504.8, respectively.

Is the method fit for purpose if the linear calibration in Figure 5-2a is used?

Figure 5-2a



Example: Testing Accuracy (2 of 3)

Solution: To be fit for purpose, the percent recovery for spiked blanks must be within 98 to 102% of the concentration added. For 1 to 10 mg/L quinine, the linear least-squares equation is $y = 92.4x + 72.9$. The quinine concentration corresponding to a signal of 194.0 is

$$x = \frac{y - 72.9}{92.4} = \frac{194.0 - 72.9}{92.4} = 1.31 \text{ mg/L}$$

The spike recovery for the 1.50-mg/L spike is

$$\% \text{ recovery} = \frac{C_{\text{spiked sample}} - C_{\text{unspiked blank}}}{C_{\text{added}}} \times 100 = \frac{1.31 - 0}{1.50} \times 100 = 87.3\%$$

For the 4.50-mg/L spike, the calculated concentration is 4.67 mg/L and its recovery is 103.8%. Spike recoveries differ from 100% by more than ±.2%. The method lacks the accuracy necessary to be fit for purpose.

Example: Testing Accuracy (3 of 3)

Test Yourself: If a blank spiked to 3.00 mg/L quinine gives a signal of 354.4, what is the predicted quinine concentration and percent recovery?

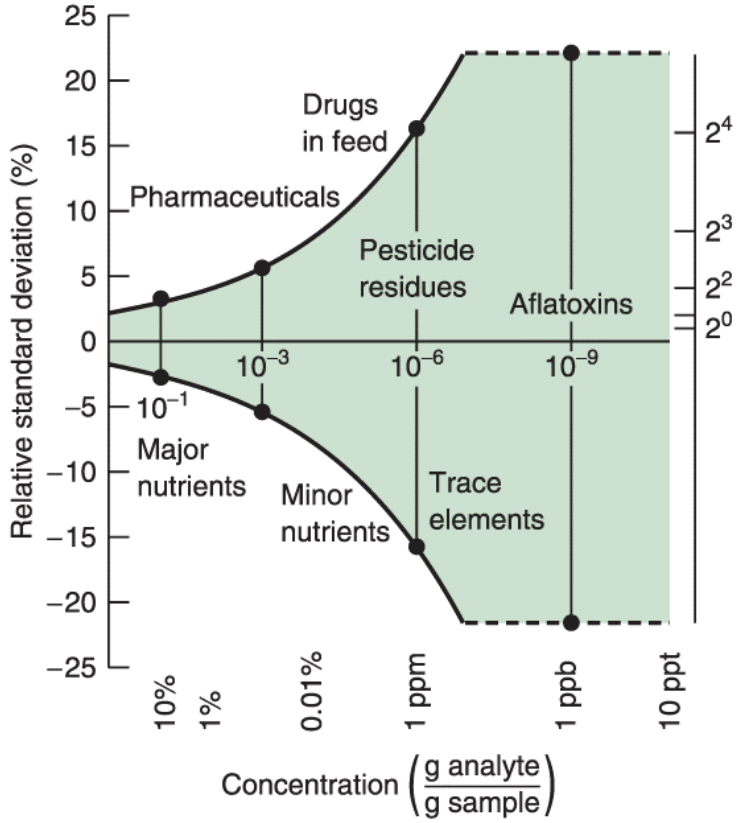
Types of Precision

- **Instrument precision**: same quantity of one sample is repeatedly introduced into an instrument
- **Intra-assay precision**: analysis of aliquots of a homogeneous material several times by one person in one day with the same equipment
- **Intermediate precision (ruggedness)**: assay is performed by different people on different instruments on different days in the same lab
- **Interlaboratory precision**: aliquots of the same sample analyzed by different people in different laboratories

Box 5-3 The Horwitz Trumpet: Variation in Interlaboratory Precision (1 of 2)

Interlaboratory tests are routinely used to validate new procedures (especially if intended for regulatory use).

- Relative standard deviations of interlaboratory results as a function of sample concentration (g analyte/g sample)
- Shaded region has been referred to as Horwitz trumpet because of the way it flares open



Harris/Lucy, *Quantitative Chemical Analysis*, 10e,

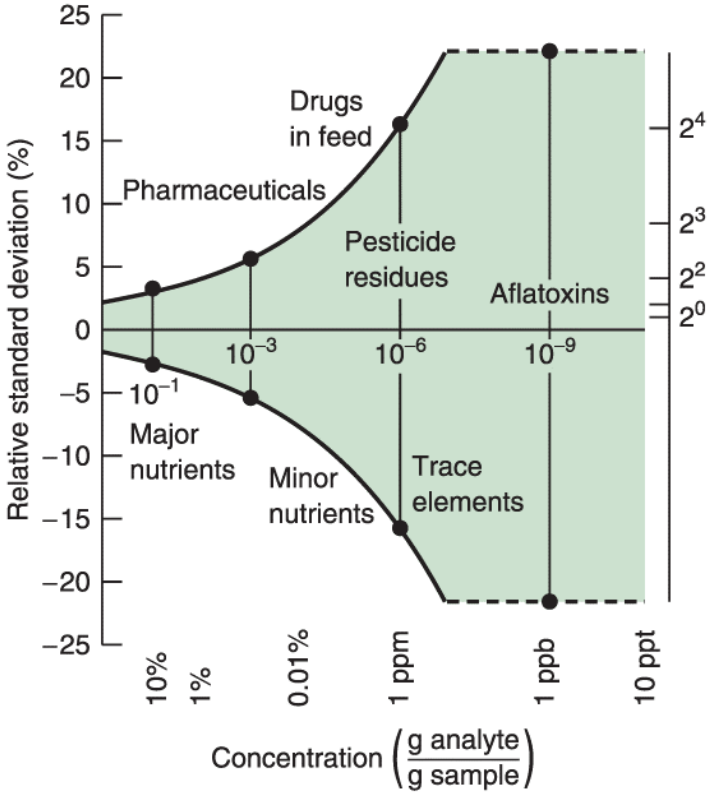
Box 5-3 The Horwitz Trumpet: Variation in Interlaboratory Precision (2 of 2)

Relative standard deviation of mean values reported by different laboratories increased as analyte concentration decreased down to 100 ppb.

$$\textit{Horwitz curve}: \ \mathrm{RSD}(\%) \approx 2^{(1-0.5 \log C)}, \text{ for} \geq 10^{-7} \text{ g/g}$$

For analyte concentrations <100 ppb, the relative standard deviation was constant at about 22%.

$$\textit{Thompson plateau}: \ \mathrm{RSD}(\%) \approx 22\%, \text{ for} \geq 10^{-7} \text{ to } 10^{-11} \text{ g/g}$$



Harris/Lucy, *Quantitative Chemical Analysis*, 10e,